TRAVEL FORECASTING MODELS FOR INDIAN URBAN STRUCTURE

by

BOMMAREDDY SATYANARAYANA REDDY

DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY, KANPUR

OCTOBER, 1984

TRAVEL FORECASTING MODELS FOR INDIAN URBAN STRUCTURE

A Thesis Submitted
In Partial Fulfilment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

by

BOMMAREDDY SATYANARAYANA REDDY

to the

DEPARTMENT OF CIVIL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY, KANPUR

OCTOBER, 1984

DEDICATED

TO

MY PARENTS

CERTIFICATE

This is to certify that the thesis 'Travel Forecasting Models for Indian Urban Structure' submitted by Shri Bommareddy Satyanarayana Reddy in partial fulfilment of the requirements for the degree of Doctor of Philosophy of the Indian Institute of Technology, Kanpur, is a record of bonafide research work carried out by him under my supervision and guidance. The work embodied in this thesis has not been submitted elsewhere for the award of a degree.

October 5, 1984

(B.R. MARWAH)
Assistant Professor
Department of Civil Engineering
Indian Institute of Technology, Kanpur

ACKNOWLEDGEMENT

The author is deeply indebted to Dr. B.R. Marwah for his invaluable guidance, encouragement and supervision throughout the course of research work and preparation of manuscript.

The author is thankful to the authorities of the Maulana Azad College of Technology, Bhopal for sponsoring him for doctoral studies under Quality Improvement Programme and to the authorities of the Indian Institute of Technology, Kanpur for providing him the necessary facilities.

The author is grateful to the authorities of Town and Country Planning Department of the State of Uttar Pradesh for providing the household survey data and to Shri G.K. Vasista for his immense care in punching the data.

The help rendered by Shri V.V.S. Ramam during the initial stages of this study is gratefully acknowledged. Valuable assistance rendered by Shri K.V. Jayakumar and Shri V. Sri Kant in proof reading is very much appreciated.

Thanks are due to Shri G.S. Trivedi for his excellent typing, to Shri J.C. Verma for tracing the drawings, and to Shri R.S. Dwivedi for duplicating.

Finally, the author would like to thank his wife, Padmavathi, and children, Raajita and Saujanya for their patience and understanding during the course of this study.

B.S.N. Reddy

CONTENTS

		PAGE
LIST OF TABLES		vii
LIST OF FIGURES		ix
LIST OF SYMBOLS AND ABBREVIATIONS		xiv
LIST OF APPENDICES		xx
SYNOPSIS		xxi
1.	INTRODUCTION	1
	1.1 Background	1
	1.2 Statement of the Problem	3
	1.3 Solution Methodology	6
	1.4 Organisation of the Report	9
2.	TRIP PRODUCTION MODELS	11
	2.1 Introduction	11
	2.1.1 Zonal Regression Models	11
	2.1.2 Category Analysis	13
	2.1.3 Household Regression Models	14
	2.1.4 Individual/Person Trip Models	15
	2.1.5 Scope of Trip Productions in the Study	17
	2.2 Multiple Regression Analysis	17
	2.2.1 Assumptions	17
	2.2.2 Model Formulation	18
	2.2.3 Selection of Best Set of Independent Variables	22
	2.3 Data Collection for Case Study	25
	2.3.1 Background	25
	2.3.2 Household Survey	28
	2.3.3 Travel Time Survey	33
	2.4 Trip Production Models for Kanpur City	34
	2.4.1 Introduction	34
	2.4.2 Zonal Regression Models	35
	2.4.3 Category Analysis	49
	2.4.4 Household Regression Models	62

PAGE

3. MODE CHOICE DECISION MODELS 74

3.1 Introduction 74
3.2 Decision Theory 75
3.2.1 Concept of Utility 77
3.2.2 Methods of Selecting a Best Action 77
3.3 Mode Choice Models for Kanpur 79
3.3.1 Introduction 79
3.3.2 Maximin Model 80
3.3.3 Maximax Model 81
3.3.4 Predominant-Attribute Model 85
3.3.5 Model Inputs 86
3.3.6 Model Outputs 91

4. SIMULATION MODEL FOR COMBINED MODAL SPLIT AND TRIP DISTRIBUTION 98

4.1 Disaggregate Models 98
4.2 Choice of Simulation for Decision Making Process 101
4.3 Concept of Status Groups 103
4.4 Decision Process Inputs 105
4.4.1 Status Group of Individual 105
4.4.2 Distribution of Employment/ Education Opportunities 107
4.4.3 Decision Profiles 109
4.5 Model Attributes 117
4.6 Model Structure 121
4.6.1 Overview 121
4.6.2 Status Group Submodel 127
4.6.3 Mode Choice Submodel 130
4.6.4 Distribution Submodel 130
4.7 Model Calibration 132
4.8 Model Validation 136
4.8.1 General 136
4.8.2 Strategy for Validation 137
4.8.3 Status Group Submodel 139
4.8.4 Mode Choice Submodel 148
4.8.5 Distribution Submodel 159
4.8.6 Conclusions 180

5. PROGRAMMING MODEL FOR COMBINED TRIP DISTRIBUTION AND ASSIGNMENT 181

5.1 Introduction 181
5.2 Review of Literature 182
5.2.1 Trip Distribution Model 182
5.2.2 Traffic Assignment Model 184
5.2.3 Combined Trip Distribution and Assignment Model 186

PAGE

5.3 Model Development 189
5.3.1 Trip Distribution Problem 190
5.3.1.1 Choice of Entropy Constraint 193
5.3.1.2 Calibration of Gravity Model Coefficients 194
5.3.1.3 Solution of the Trip Distribution Problem 200
5.3.2 Combined Distribution and Assignment Problem 202
5.3.2.1 Problem Formulation 202
5.3.2.2 Solution of the Combined Distribution and Assignment Problem 207
5.4 Model Application 209
5.4.1 Strategy for Model Application 209
5.4.1.1 Link Capacities 211
5.4.1.2 Selection of Significant O-D Pairs 213
5.4.2 Gravity Submodel 215
5.4.3 Combined Distribution and Assignment Model 226
5.4.3.1 Selection of Significant Node Pairs 226
5.4.3.2 Analysis of Model Outputs 231

6. SUMMARY AND CONCLUSIONS 240

6.1 Summary 240
6.2 Conclusions 247

REFERENCES 251

APPENDICES I TO II 260

LIST OF TABLES

TABLE No.	TITLE	PAGE
2.1	Distribution of Mode Choice for Sample Household Survey Data	33
2.2	Relations for Zonal Trip Productions	37
2.3	Statistical Analysis for Zonal Trip Productions	39
2.4	Relations for Zonal Work Trip Productions by Individual Modes	41
2.5	Relations for Zonal Education Trip Productions by Individual Modes	44
2.6	Statistical Analysis of Zonal Trip Productions by Individual Modes	47
2.7	Household Daily Trip Production Rates for Family Size-Income Level Category	53
2.8	Household Daily Trip Production Rates for Family Size-Vehicle Ownership Category	55
2.9	Household Daily Trip Production Rates for Income Level-Vehicle Ownership Category	56
2.10	Household Daily Trip Rate by Purpose for Family Size-Income Level Category	59
2.11	Household Daily Trip Production Rate by Purpose for Family Size-Vehicle Ownership Category	60
2.12	Household Daily Trip Production Rate by Purpose for Income Level-Vehicle Ownership Category	61
2.13	Comparison of Observed and Model Zonal Trip Productions	63
2.14	Relations for Household Work and Education Trip Productions	68
2.15	Statistical Analysis of Household Work and Education Trip Productions	73

TABLE No.	TITLE	PAGE
3.1	Observed and Model Trips by Different Modes of Travel for Work Trips	92
3.2	Observed and Model Trips by Different Modes of Travel for Education Trips	94
3.3	Observed and Model Trips When Motor Vehicular Modes are Considered for Work Trips	96
3.4	Observed and Model Trips When Motor Vehicular Modes are Considered for Education Trips	97
4.1	Distribution of Income Levels in Cluster of Zones Between Cordons	140
4.2	Distribution of Educational Levels in Cluster of Zones Between Cordons	141
4.3	Comparison of Trips in Cluster of Zones Between Cordons	144
4.4	Distribution of Trip Productions for each Status Group	146
4.5	Analysis of Trips by Different Modes of Travel	150
4.6	Work Trip Mode Choice for Each Status Group	155
4.7	Education Trip Mode Choice for Each Status Group	157
5.1	Final Calibrated Coefficient (β_0) for the Relationship $f_{ij}=A\ \mathrm{ExP}(-\beta_0 C_{ij})$	224
5.2	Maximum and Planned Entropy Values	227
5.3	Objective Function for Different Modes	230
5.4	Entropy Values for Different Modes	232

LIST OF FIGURES

FIGURE No.	TITLE	PAGE
2.1	Kanpur City Road Network	27
2.2	Zoning of Kanpur City	30
2.3	Trip Production Rates for Family Size-Income Level/Vehicle Ownership Categories	54
2.4	Trip Production Rates for Income Level-Vehicle Ownership Category	57
3.1	Maximin Model	82
3.2	Maximax Model	84
3.3	Predominant-Attribute Model	87
4.1	Status Group of an Individual	106
4.2	Distributions of Employment/Education Opportunities	108
4.3	Procedure for Determining Decision Profiles	110
4.4	Decision Profile for Travel Modes	112
4.5	Decision Profile for Employment/Education Opportunities	113
4.6	Decision Profile of Travel Times for Status Group 8	115
4.7	Decision Profile of Travel Times for Mode 5 (Bus)	116
4.8	Decision Profile of Travel Costs for Status Group 8	118
4.9	Decision Profile of Travel Costs for Mode 5 (Bus)	119
4.10	Structure of the Model	122
4.11	Interaction of Submodels	128
4.12	Decision Profile for Income Levels in Central Cordon	134
4.13	Decision Profile of Status Groups	135

FIGURE No.	TITLE	PAGE
4.14	Cumulative Frequency Distributions of Income/Educational Levels Within Cordon 1 (Work Trips)	142
4.15	Frequency Distribution of Status Groups for Work Trips	145
4.16	Frequency Distribution of Status Groups for Education Trips	145
4.17	Comparison of Observed and Simulated Trips by Different Status Groups	147
4.18	Cumulative Frequency Distribution of Trips by Different Modes (Work Trips)	149
4.19	Cumulative Frequency Distribution of Trips by Different Modes (Education Trips)	149
4.20	Comparison of Observed and Simulated Trips by Different Status Groups (Bicycle Trips)	152
4.21	Comparison of Observed and Simulated Trips by Different Status Groups (Motorcycle/scooter Trips)	153
4.22	Comparison of Observed and Simulated Trips by Different Status Groups (Cycle rikshaw Trips)	154
4.23	Comparison of Observed and Simulated Work Trips Between Different O-D Pairs	160
4.24	Comparison of Observed and Simulated Education Trips Between Different O-D Pairs	161
4.25	Comparison of Observed and Simulated Trip Attractions to Various Destination Groups	163
4.26	Cumulative Frequency Distribution of Trips by Different Destination Groups (Work Trips)	164
4.27	Cumulative Frequency Distribution of Trips by Different Destination Groups (Education Trips)	164

FIGURE No.	TITLE	PAGE
4.28	Frequency Distribution of Trips by Different Destination Groups for Status Group 1 (Work Trips)	165
4.29	Frequency Distribution of Trips by Different Destination Groups for Status Group 1 (Education Trips)	165
4.30	Comparison of Observed and Simulated Trips by Different Travel Time Groups	166
4.31	Cumulative Frequency Distribution of Trips by Different Travel Time Groups (Work Trips)	167
4.32	Cumulative Frequency Distribution of Trips by Different Travel Time Groups (Education Trips)	167
4.33	Frequency Distribution of Trips by Different Travel Time Groups for Status Group 1 (Work Trips)	168
4.34	Frequency Distribution of Trips by Different Travel Time Groups for Status Group 1 (Education Trips)	168
4.35	Frequency Distribution of Travel Times for Different Status Groups (Work Trips)	170
4.36	Frequency Distribution of Travel Times for Different Status Groups (Education Trips)	171
4.37	Comparison of Observed and Simulated Trips by Different Travel Cost Groups	172
4.38	Cumulative Frequency Distribution of Trips by Different Travel Cost Groups (Work Trips)	173
4.39	Cumulative Frequency Distribution of Trips by Different Travel Cost Groups (Education Trips)	173
4.40	Frequency Distribution of Trips by Different Travel Cost Groups for Status Group 1 (Work Trips)	174
4.41	Frequency Distribution of Trips by Different Travel Cost Groups for Status Group 1 (Education Trips)	174

FIGURE No.	TITLE	PAGE
4.42	Frequency Distribution of Travel Costs for Different Status Groups (Work Trips)	176
4.43	Frequency Distribution of Travel Costs for Different Status Groups (Education Trips)	177
4.44	Frequency Distribution of Travel Time Groups (Work Trips)	178
4.45	Frequency Distribution of Travel Time Groups (Education Trips)	178
4.46	Frequency Distribution of Travel Cost Groups (Work Trips)	179
4.47	Frequency Distribution of Travel Cost Groups (Education Trips)	179
5.1	Network of Centroids and Interzonal Links	212
5.2	Friction Factor-Travel Time Relation (1st Iteration)	217
5.3	Friction Factor-Travel Time Relation (3rd Iteration)	217
5.4	Friction Factor-Travel Time Relation (5th Iteration)	218
5.5	Friction Factor-Travel Time Relation (7th Iteration)	218
5.6	Friction Factor-Travel Time Relation (9th Iteration)	219
5.7	Friction Factor-Travel Time Relation (12th Iteration)	219
5.8	Travel Time Frequency Distribution (1st and 3rd Iteration)	220
5.9	Travel Time Frequency Distribution (5th and 7th Iteration)	221
5.10	Travel Time Frequency Distribution (9th Iteration)	222
5.11	Travel Time Frequency Distribution (12th Iteration)	223
5.12	Frequency Distribution of Different Travel Times For Bus (Work Trips)	233
5.13	Frequency Distribution of Different Travel Times for Car (Work Trips)	233

FIGURE No.	TITLE	PAGE
5.14	Frequency Distribution of Different Travel Times for Motorcycle/Scooter (Work Trips)	234
5.15	Frequency Distribution of Different Travel Times for Bus (Education Trips)	234
5.16	Comparison of Observed and Simulated Interzonal Transfers for Bus	235
5.17	Comparison of Observed and Simulated Interzonal Transfers for Car	236
5.18	Comparison of Observed and Simulated Interzonal Transfers for Motorcycle/Scooter	237
5.19	Flows Along the Interzonal Links of the Network	239

LIST OF SYMBOLS AND ABBREVIATIONS

A	Calibration parameter
a_o	A constant included to represent that portion of the value of Y
a_i	i = (1,2,...,n) regression constants
a_j	Number of trips per unit time ending in zone j
a_{ij}^{pq}	Link incidence matrix
CBD	Central Business District
C_p	Total square error for p parameters
c	Cluster of zones between cordons
c_{ij}	Unit generalized cost (or travel time) from origin i destination j.
c_p	Cost per unit flow on link p
DEC-1090	DEC 1090 computer system
DFP	Davidon-Fletcher-Powell
D_j	A constant
d	Destination zone group
E	Planned entropy
E_{max}	Maximum entropy
E_{min}	Minimum entropy
E_O	Existing entropy
EDL_I	Education level of Individual I
EM_i	Number of education trips produced in the zone by ith mode (i=1,2,....,6)
ExP	Exponential
F	Objective function
FS_k	Number of households having family size of kth group (k=1,2,....,4)
f_i	Value of friction factor for i th interval as used in the calibration just completed

f_i'	Value of friction factor for the ith interval to be used in the next iteration
f_{ij}	Friction factor
GM_i	Percentage of total trips occuring for ith travel time/cost interval obtained from the model
g	Constraint
$[H]$	A positive definite symmetric matrix
I	Individual
IG_l	Number of households having income level of lth group ($l = 1,2,\ldots,4$)
INC_I	Income group of individual I
$INCOBS_c$	Income/educational levels in cluster of zones c between cordons
i	Zone of origin
j	Alternative with minimum utility for an attribute (Chapter 3)
j	Destination zone (Chapter 4,5)
k	Number of independent variables entered in the regression equation
L_p	Link p ($p = 1,2,\ldots, NL$)
M	Number of design variables
M_1	Number of inequality constraints
MSE	Mean square due to residuals
MSE_p	Mean square error for p parameters
MSR	Mean square due to regression
m	Mode of travel (Chapter 4)
m	Number of zones in which trip begins (Chapter 5)
m_1	Number of undirected links
N	Total number of observations (Chapter 2)
N	Number of constraints (Chapter 5)

N_1	Number of equality constraints
NAT_j	Total number of attractions to zone j
$NATOBS_i$	Number of employment/education opportunities for zone i
$NATSIM_j$	Number of trip attractions to zone j out of the already simulated trips
$NAT(ZONES_{d,n})$	Total number of trip attractions in the nth zone of the destination group d
$NEOOBS_{s,d}$	Number of employment/education opportunities of status group s for destination group d
NL	Total number of links along the path
$NSOBS_s$	Number of individuals in status group s
$NSSIM_s$	Number of individuals simulated in status group s out of the already simulated individuals
NT	Number of travel time/cost intervals
NT_i	Number of trips produced in zone i
$NTCOBS_{s,tc}$	Number of trips by different travel cost groups tc for status group s
$NTMOBS_{s,m}$	Number of trips by different modes of travel m for each status group s
$NTMSIM_{s,m}$	Number of trips by mode m for status group s out of already simulated trips
$NTMSIM(SONES_{d,n})$	Number of trips attractions in the nth zone of the destination group d out of the already trips
$NTTOBS_{s,tg}$	Number of trips by different travel time groups tg for status group s
n	Number of independent variables (Chapter 2)
n	Number of zones in which trip ends (Chapter 5)
n_a	Number of attributes of interest in mode choice
n_m	Number of alternative modes of travel
n_1	Number of O-D pairs

O_i	A constant
O-D	Origin-Destination
OD_i	Percentage of total trips occuring for ith travel time/cost interval
P_i	Number of trips per unit time begining in zone i
R	Multiple correlation coefficient
R^2	Coefficient of multiple determination
R_p^2	Coefficient of multiple determination for p parameters
r	Penalty parameter
$\bar{S}$	Search direction
SEE	Standard error of estimate
SEEP	Percentage of the mean observed value of dependent variable Y
SSR	Regression sum of squares
s	Status group
T	Total number of trips
TBC	Number of bicycles in a zone
TC	Travel cost group
$TC(m,i,ZONES_{d,n})$	Travel cost for nth zone in the destination group d from zone i by mode m
TC_A	Total cost due to assignment
$TCOS_{tc}$	Travel cost for tcth group
$TCOS_{tc-1}$	Travel cost for (tc-1)th group
TEM	Total education trip productions with all modes combined (number of education trips produced in the zone)
TGM_i	Trips observed from the model for ith travel time interval
TIC	Total income level of the zone

TL_n	Number of households having trip length of nth group (n=1,2,....,5)
TMC	Number of two wheelers (motorcycle and scooters) in a zone
TN_{ij}	Total number of paths between i and j
TNP	Total number of persons in the zone
TOB_i	Observed trips for ith travel time interval
$TT(m,i,ZONES_{d,n})$	Travel time for nth zone in the destination group d from origin i by mode m
TTL	Total distance travelled by the trip makers of the zone (zonal trip length)
TTM_{tg}	Travel time for tg th group
TTM_{tg-1}	Travel time for (tg-1)th group
TWM	Total work trip productions with all modes combined (number of work trips produced in the zone)
t_{ij}	Number of trips from origin i to destination j
tc	Travel cost group
tg	Travel time group
$U_{i,j}$	Minimum utility of ith attribute that occur for jth alternative
$U_{i,k}$	Utility of jth attribute for kth alternative
$U_{i,l}$	Maximum utility of ith attribute that that occur for lth alternative
$U'_{i,k}$	Relative utility value for an attribute i of alternative k
$U'_{p,j}$	Relative utility of the most predominant attribute p, for the jth alternative having maximum utility
$U'_{p,k}$	Relative utility of the most predominant attribute p for each of the alternatives k

$U_{s,j}$	Utility of sth attribute for jth alternative
$U_{s,l}$	Utility of sth attribute for lth alternative
V_m	Number of households having vehicle ownership of mth group (m=1,2,...,8)
WM_i	Number of work trips produced in the zone by ith mode (i=1,2,....,6)
X_i	Explanatory or independent variables (i=1,2,..,n)
Y	Response or dependent variable
Y'	Value of Y calculated from the regression equation
$\bar{Y}$	Mean value of Y
z	Nonnegative constant
α	Step length
α^*	Minimizing step length
$\bar{\alpha}$	Vector of design variables
β	Calibration parameter
$\emptyset$	Modified objective function
e	Error term
λ_k	A variable

LIST OF APPENDICES

APPENDIX No.	TITLE	PAGE
I	Details of Traffic Zones with Cordon Numbers for Kanpur City	260
II	Alternative Paths Between Important O-D Pairs	263

SYNOPSIS

TRAVEL FORECASTING MODELS FOR INDIAN URBAN STRUCTURE

(A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy by Bommareddy Satyanarayana Redd to the Department of Civil Engineering, Indian Institute of Technology, Kanpur, India, October, 1984)

Process of urbanization in developing countries is not simply a repetition of what occurred in the developed world, rather it is occurring much faster and under different economic and technological conditions. Large size developing cities generally suffer more from traffic congestion, road accidents, environmental pollution, overcrowding of public transport, and poor conditions for pedestrians.

In the last three decades planning of urban transport has changed dramatically and a large number of models have been developed for various stages of the travel forecasting process. These models depend on a high level of car ownership and a relatively homogeneous distribution of land use at low density. The cities in developing countries have lower income level, multiple modes of travel, heterogeneous distribution of land use at high density and poor transport facilities. The factors that generate trips and which determine mode choice among the low income groups are not well understood and cannot be deduced from experience in developed countries. A study of Indian social structure in relation to travel habits is necessary to develop trip generation models.

The aim of this study is to develop system of models appropriate for Indian urban structure, which given a knowledge of land use, household characteristics, and available transport facilities could estimate accurately the intricate movements of traffic on the system by mode. The models evolved in the study include: trip production models; mode choice decision models based on non-compensatory choice processes; simulation of modal split and trip distribution as an interconnected decision process; and programming model for combined trip distribution and assignment. These models could be used to show what would happen if transport facilities are changed and if over time, the land use and household characteristics are altered. The efficient methodologies for travel demand models can be evolved only when carefully programmed basic data related to the existing situation of the system are available. The models evolved are calibrated and validated for the city of Kanpur, the premier industrial metropolis of Northern India. The city is divided into 56 traffic zones and the models are calibrated for the base year household travel survey data.

Trip Production Models:

Three trip production models, namely, zonal regression category analysis, and household regression models are constructed based on the household survey data. For the selection of best set of explanatory variables in the zonal and household regression models, forward stepwise regression method is used. The explanatory variables considered for zonal regression are: number of persons in the zone, total income level for the zone, number of vehicles

of different types in a zone, and zonal trip length. Zonal trip production relationships are obtained for the work and education trips by modes. Considering the type of household data available and with the aim to conserve homogeneity within each type of household, category analysis is used for productions of work and education trips. The household is classified in three different ways such as family size-income level, family size-vehicle ownership, and income level-vehicle ownership combinations. In household regression approach the analysis is directly applied to the basic household data, and it operates on the whole of the variability which exists in the data at that level. This approach is a combination of better features of zonal regression and category analysis. The choice of explanatory variables are guided by the results of zonal regression and household classifications. The explanatory variables that have been considered are: family size, income level, vehicle ownership, and trip length. These variables are further stratified into 21 subgroups. Because of correlation between the variables, number of explanatory variables that are found to enter the regression equations range between seven and fifteen.

Mode Choice Decision Models

Mode choice decision models of non-compensatory structures are made on the basis of attribute by attribute comparison of various alternatives. The levels of satisfaction associated with each of the attributes perceived by the person in the decision process are evaluated. Based on the concept of decision theory maximin, maximax, and predominant-attribute models are formulated.

The choice of attributes involved in the decision making process is of paramount importance. In a typical metropolitan city of India having number of alternative modes, the decision maker has to access his/her own socio-economic characteristics in relation to the characteristics of available alternative modes. The attributes considered in this study are: family size, income level, vehicle ownership, and trip length.

The maximin/maximax strategy calls for the identification of least/highest satisfactory alternative of each attribute and the selection of the alternative which yields the highest minimum/maximum level of satisfaction. Predominant-attribute model calls for identification of the most predominant attribute according to the generic importance. Alternative having the highest level of satisfaction for the predominant attribute is selected. The results obtained from maximin and maximax models compare well with the observed data, but the results obtained from the predominant-attribute model significantly vary from the observed values. As such the most predominant-attribute model as formulated does not appear to be suitable for the city of Kanpur. This may be attributed to the fact that the utility value for each attribute is taken as the proportion of trips for that alternative, or it may be that the mode choice decision process is not affected only by the most predominant attribute. It is desirable to test this model for some other cities also to identify the appropriate causes for differences.

Simulation of Modal Split and Trip Distribution

Different designs for disaggregate models have been suggested to solve complex interaction processes, but the implementation has been highly hampered by the methodological difficulties. Simulation of the complex decision process at the individual level is attempted in this study. The model simulates modal split and trip distribution as an interconnected decision process. The model essentially consists of status group,mode choice, and distribution submodels. Individuals with similar constellations of activity constraints are grouped into status groups. The individual decision process which is stochastically determined from the decision profile of his/her status group,identifies the mode choice and trip destination. The decision profiles which act as the major constraints on travel behaviour include: distributions of productions and attractions; distribution of zonal income levels; distributions of travel mode, travel time, travel cost, and employment / education opportunities for each of the status group; and travel time and travel cost distributions by mode. The components of the model are activated in a specific order and with certain specific feedbacks. The model is so structured that the order of activities of the different components can be easily changed as per the framework of the study.

The simulation procedure first creates an individual. From the origin of the individual and income/education level,the status group is identified. The mode of travel is identified from the decision profile of travel generated for

the relevant status group. The mode selected remains fixed during the decision cycle of the individual's destination choice. If it so happens that the destination could not be assigned, then a new attempt on the selection of the mode is made. The travel time range for the individual is selected from the travel time decision profile of the relevant status group. Using the decision profile of employment/education opportunities ,a set of destination zones that have the potential to attract, and can be reached within the assigned travel time range are identified. When more than one destination zones are identified, a travel cost range is randomly drawn from the appropriate decision profile. Out of the already identified set of destinations,only those which can be reached within the selected travel cost range are sorted, and the one with minimum travel cost is selected provided it has already not been assigned the desired trip ends.

The empirical validation of the simulation model is attempted by comparing the model outputs with the observed household survey data. Each of the submodels is validated by appropriate measures of effectiveness, so as to ensure that they have been assembled into a realistic structure. The results of the validation procedure indicate the capability of the model to simulate the decision process of the individual in a complex multimodal situation. The algorithm is quite sensitive and also flexible enough to incorporate alternative decision processes.

Programming Model for Combined Trip Distribution and Assignment

An attempt is also made in this study to derive models that combine trip distribution and traffic assignment into a single stage. This model optimizes the system by minimizing the total cost of the system as a whole. The main components of the model are entropy type trip distribution submodel for each mode, and a system optimized assignment submodel. The combined model is solved by using the decomposition principle. All the constraints of the gravity submodel of the distribution problem are linear except the entropy constraint which is convex in the region defined by linear constraints. The gravity submodel is calibrated by making use of Davidon-Fletcher-Powell (DFP) method of constraint optimization.

The gravity submodel through an iterative process evolves the distribution matrix for each mode within the framework of the constraints of the given trip generation data, the generalized cost, and existing extent of entropy. The distribution matrix thus obtained is the input to the assignment submodel. The assignment submodel minimizes the total travel cost through system optimization for each mode considering the link costs over the network. The dual costs obtained from the assignment submodel replace the generalized costs in gravity submodel as feedback so as to get the revised trip matrices. These two submodels stop interacting when no further change in objective function is possible.

The assignment model requires alternative paths between each O-D pair as an input. In this study only two alternative

paths between each O-D pair are considered which include the shortest travel time path and the one with next higher travel time. The capacity of each of the links also act as a constraint while assigning the trips. The model solution involves large magnitude of core capacity because of the requirements of work vector, the coefficient matrices and resource vector inbuilt in the linear programming. The available DEC-1090 system is not capable of handling the problems of this magnitude. Hence it is decided to consider atmost 100 significant O-D pairs for assignment by the model. A strategy is planned for identifying O-D pairs having significant number of interzonal transfers. The assignment for these significant pairs is done by the model. Transfers between other O-D pairs are assigned to the shortest paths. The comparisons of the observed and simulated frequencies of trips for each of the travel time groups indicate a close agreement. The simulated interzonal transfer values are also close to the observed ones. The link flows as obtained from the model are compared with link flows when trips are assigned only to shortest paths. The results indicate that even for the city of Kanpur, where not many alternative paths are available, flows on some of the links do get affected while optimizing the system costs.

The greatest contribution of the study is that the algorithms evolved are practical for problems encountered in cities of developing countries. Further the validity of the various models is clearly established. To enhance the achievements of the study, the models may be further refined and tested for a number of cities.

1 INTRODUCTION

1.1 BACKGROUND

One of the most daunting problems faced by the cities is of urban transport. A solution to the transport problem must consist of a land use pattern, a transport system and a set of management policies that together bring demand and supply into balance, both for transport services and for those other amenities affected by transport or which compete with transport for space. A transport plan is intended to tackle the problem as it presently exists and is expected to develop in the future.

In the last three decades, planning of urban transport has changed dramatically. Until the mid-1950s, it consisted of estimating traffic demand by means of simple growth rates. The mid-1950s saw the advent of large scale urban transport studies. Scientific planning, through which traffic demands were related to land use and modal choice, was introduced. The overwhelming concern of these studies was for the road user, its costs, travel times, and safety. The methods were pioneered in Detroit and Chicago (Detroit, 1955; Chicago, 1959, 1960, and 1962), where the object was almost exclusively to provide roads for private transport with little attention to public transport.

In the 1960s transport planning went through an evolutionary step known as 'big plan' phase in which massive

studies aimed to produce continuous comprehensive long-term plans for land use and transport in considerable detail (Thomson, 1983). The essence of this new method lay in recognition of interaction between land use and transport, and between different parts of the transport system. Shortly thereafter mass transit also became a focus of concern. The later years saw the advent of increased concern for social and environmental aspects of transportation and for involvement of travellers in the decision making process. The major thrust of the transportation planning study involves the determination of demand for travel. This is followed with an attempt to find those presently available or anticipated future transport system which could meet the established travel demands in an acceptable manner. The travel demand prediction usually is accomplished by four step procedure which uses predicted zonal land use values as well as future transportation system performance characteristics. The steps are: trip generation, trip distribution, mode choice, and traffic assignment.

The 'big Plan' approach produced some useful results that led to a better understanding of the problem, but it also produced inadequate results, some of which were inherent in the method itself. The models failed to reflect realistically the interactions between land use and transport, response of public transport passengers to changes in service, and could not produce adequately the phenomena of congestion. The plan took a long time to produce, tended to ignore budget constraints and neglected the role of demand management as an alternative

to investment. From time to time voices have been raised demanding that the 'big plan' approach models should be simplified. These voices are a reaction to the cost, duration, intelligibility and applicability of the results of the 'big plan' models.

The oil crisis of 1973 and the following world-wide recession tended to change the attitude toward transport planning. There was a feeling that the 'big plans' were a waste of time and money, attempting the impossible by being too ambitious, without doing what might be possible. This critical reaction led to the 'incremental improvement' phase in which planning was confined to immediate action and short-term improvements with strong emphasis on low cost measures. This approach leads rapidly to results which can be seen to be good at least for a time, but the shortcomings are also fairly clear. Transport improvements which appear obviously desirable at a time may be ultimately counter productive. Moreover short-term plans are basically short-sighted and fail to give due weight to long-term investments.

1.2 STATEMENT OF THE PROBLEM

The urban transport problem is fundamentally similar in all large size cities. The basic causes are the same and so are many of the consequences. There are however differences of degree between the cities of the developed and developing countries The cities in the developed world have a much greater level of

vehicle ownership and also a higher rate of trip generation than in the developing cities, still it is the developing cities that generally suffer more from traffic congestion, road accidents, environmental pollution, overcrowding of public transport, and poor conditions for pedestrians and man driven vehicles. Besides these problems it is important to determine if the right journeys are made in the right way, since the real test of the system is whether people can obtain access to their activities without undue expenditure of time, money, and effort. Accessibility not mobility is what matters.

It may not be true to think that the process of urbanization in developing countries is only a repetition of what occurred earlier in the developed world and could, if desired, follow a similar course. The process is occurring much faster and under different economic and technological conditions. Large size cities in a developing country like India are growing rapidly to a size that has no parallel in the developed world. Failure to provide adequate transport facilities greatly increases the trip durations and costs for passengers, lowering productive efficiency and placing a particularly heavy burden on poorer groups of the population living in peripheral and other areas of very limited access.

Travel demand forecasting is a highly important part of the urban transport planning process. Its purpose is to develop tools for predicting how the transport situation will develop in the future under alternative assumptions about the factors determining transport demand on the one hand, and

changes in transport facilities, on the other. A large number of simple and sophisticated models have been developed for various stages of the travel forecasting process. These models depend on a high level of car ownership, a relatively homogeneous distribution of land use at low density, require lot more data, technical staff and regulatory capacity. These conditions are significantly different than those in the cities of developing countries. In large size cities of India with high uncertainty concerning future growth, income and land usage, the models as produced in developed contries are not appropriate. The medium to large size cities in India have generally lower income level; multiple modes of travel like bicycle, cycle rikshaw, public transport, private motor vehicles; heterogeneous distribution of land use at high density; and poor transport facilities.

Very little work has been done to develop appropriate travel estimation models for India. There is inadequate understanding of the transport behaviour in the cities. The factors that generate trips and which determine mode choice among the low income groups are not well understood and cannot be deduced from experience in developed countries. A study of Indian social structure in relation to travel habits is necessary to develop the trip generation models.

Models are a simplification of reality. The planner must concentrate on those aspects of the complex solution that are relevant to the kind of policies he is studying. In the complex travel behaviour process for Indian cities, there is a need to design models based on sound theories and valid

statistical estimation procedures. Models cannot be employed without first investigating in detail a particular situation under study, the main reason for it being that parameters that remain stable from study area to study area have not been found. As a consequence data need to be collected in each transportation study to calibrate the models.

1.3 SOLUTION METHODOLOGY

The aim of this study is to develop system of models appropriate for Indian urban structure, which, given a knowledge of land use, household characteristics, and available transport facilities could estimate accurately the intricate movements of traffic on the system by mode. The models could be used to show what would happen if transport facilities are changed, and if over time, the land use and household characteristics are altered. The efficient methodologies for travel demand models can be evolved only when carefully programmed basic data related to the existing situation of the system are available. The models evolved in this study are calibrated and validated for the city of Kanpur, the premier industrial metropolis of Northern India.

Three trip production models, namely, zonal regression, category analysis, and household regression models are constructed based on the household survey data. Work and education trips are disaggregated into six travel modes. Zonal trip production relationships by mode are obtained both for the work and education trips. Considering the type of household data available and with the aim to conserve homogeneity

within each type of household, category analysis is used for trip productions. The household is classified in three different ways such as family size-income level; family size-vehicle ownership; and income level-vehicle ownership combinations. Household regression approach is a combination of better features of zonal regression and category analysis. The analysis is directly applied to the basic household data and it operates on the whole of the variability which exists in the data at that level.

Mode choice decision models of non-compensatory structures are made on the basis of attribute by attribute comparison of various alternatives. The level of satisfaction associated with each of the attributes perceived by the person in the decision process can be evaluated. These levels help in decision making. Based on the concept of decision theory maximin, maximax, and predominant-attribute models are formulated. The choice of attributes involved in the decision making is of paramount importance. The maximin/maximax strategy calls for the identification of least/highest satisfactory alternative of each attribute, and the selection of the alternative which yields the highest minimum/maximum level of satisfaction. Predominant-attribute model calls for identification of the most predominant attribute according to the generic importance. Alternative having the highest

level of satisfaction for the predominant attribute is selected.

Different designs for disaggregate models have been suggested to solve complex interaction processes, but the implementation has been highly hampered by the methodological difficulties. Simulation of the complex decision process at individual level is attempted in this study. The model simulates modal split and trip distribution as an interconnected decision process. The proposed model essentially consists of status group, mode choice, and distribution submodels. Individuals with similar constellations of activity constraints are grouped into status groups. The individual decision process which is stochastically determined from the decision profile of his/her status group identifies the mode choice and trip destination. The proposed algorithm is sensitive and flexible to incorporate alternative decision process.

An attempt is also made in this study to derive models that combine trip distribution and traffic assignment into a single stage. The programming model optimizes the system by minimizing the total cost of the system as a whole. The main components of the model are entropy type trip distribution submodel for each mode, and a system optimized assignment submodel. The combined model is solved by using the decomposition principle. All the constraints of the gravity submodel of the distribution problem are linear except the entropy constraint which is convex in the region defined by linear constraints. The gravity

submodel is proposed to be calibrated by Davidon-Fletcher-Powell (DFP) method of constraint optimization. The gravity submodel through an iterative process evolves the distribution matrix for each mode within the framework of the constraints of the given trip generation data, the generalized cost, and existing extent of entropy. The assignment submodel minimizes the total travel cost through system optimization for each mode, considering the link costs over the network. The dual costs obtained from the assignment submodel replace the generalized costs in gravity submodel as feedback, so as to get the revised trip matrices. These two submodels stop interacting when no further change in objective function is possible. Through the successive solutions of these two submodels, the final assigned trips can be obtained.

1.4 ORGANISATION OF THE REPORT

(i) Trip production models are reviewed. Zonal regression, category analysis, and household regression models are calibrated for Kanpur city (Chapter 2).

(ii) Mode choice decision models based on non-compensatory choice are reviewed. Maximin, maximax and predominant-attribute models based on decision theory are formulated. The results of models are analysed and a comparison is made (Chapter 3).

(iii) Disaggregate models are reviewed. A simulation model for combined modal split and trip distribution is formulated. The validation of the model and associated submodels are presented (Chapter 4).

(iv) Trip distribution, traffic assignment, and combined trip distribution and assignment models are reviewed. The programming model for combined trip distribution and assignment is formulated. Strategy for model application is explained and the results are analysed (Chapter 5).

(v) Study is summerized, conclusions are drawn and suggestions are made for future investigations (Chapter 6).

2 TRIP PRODUCTION MODELS

2.1 INTRODUCTION

Urban transportation planning process incorporates the development of a representative trip generation model from base year survey data for making forecasts of travel in subsequent years. Four methods are generally used for explaining trip generation. The first method called the zonal regression model uses the traditional approach of multiple regression analysis to explain zonal trips in terms of zonal characteristics. The main alternative in recent years has been the category analysis or more generally the cross classification of household trip rates (Wootton and Pick, 1967). In view of growing interest in household regression as a means of conserving the essential variability of household trip data, the third approach namely household trip model applies regression analysis directly to the household data (Kannel, 1972; Douglas, 1973). The fourth method called the individual/person trip model (Smith and Cleaveland, 1976) also adopts the multiple regression analysis to explain the trip productions based on person trips and characteristics. A brief review of these models is given in the following sub-sections.

2.1.1 Zonal Regression Models

These models adopt the approach of applying multiple regression analysis to explain zonal trips in terms of zonal

characteristics. This procedure has been widely employed in transportation studies. Zonal trip production models are developed using socio-economic data comparable over time periods (Ashford and Holloway, 1972). The study area is divided into a number of zones and each zone is treated as one observation. Zonal values for the dependent variables are obtained by expanding the trip information from the sample household surveys into trip totals for each zone (Downes,1976). The dependent variable is some measure of the number of zonal trip ends. Trips may be stratified by purpose at the non-home end by mode of travel and by household vehicle ownership. The independent variables considered include the number of persons, number of different types of vehicles, income level distribution, number of employed persons, and school going children in a zone. Relationships between the dependent and independent variables are obtained using standard linear multiple regression techniques. Matrices of linear correlation coefficients for all pairs of variables are examined and regressions are made on selected combination of variables, testing for statistical significance and removing or replacing variables as necessary. The statistical quality and general plausibility of the relationships and their ability to reproduce the observed quantities are all taken into account in making the final selection.

The advantage of these models are that they are able to handle enormous quantity of data, and immediate expansion of sample data for calibration against screen line and other check data. The limitations of these models are that there are practical difficulties in applying these results to other

zone systems. Further the equations do not provide an explanation of trip generation at levels lower than the zonal aggregate data on which they are based.

2.1.2 Category Analysis

It is a technique based on the household and its characteristics without entering into statistical problems of linear least squares. It is a procedure determining the average response or average value of the dependent variable for a set of categories of independent variables. Usually the average response is taken to be the average number of trips made by a particular category of analysis unit (household or zone). The categories are defined by levels of those causal variables which are supposed to influence trip generation (income level, vehicle ownership, family size etc.,).Unlike least squares regression analysis, category analysis does not require the formulation of any mathematical model (Wootton and Pick, 1967;Ashish Sen and Johnson, 1977).

The main advantage of this model over zonal regression model is that it provides an opportunity for examining the linkages between the purpose and mode of travel, person types and mobility level of households. As it is independent of zone system the mean trip rates per household type can be applied to any system of aggregation. The limitation of this model is that the subdivision of the household sample into many household types can produce uneven cell sizes making some of the mean trip rates less reliable than others. The

other drawback of this model is the difficulty of including some measure of accessibility to centres of urban activity without involving the zone system and reducing the flexibility of application.

2.1.3 Household Regression Models

Household regression is a composite approach to trip generation which combines the better features of zonal regression and household classification. It provides a linear regression equation instead of a catalogue of mean trip rates for different classes of household characteristics instead of zonal characteristics. The basic characteristics of these models are that the analysis is applied to the basic household data and operates on the whole of the variability which exists in the data. These models are developed by using the multiple regression techniques to the complete set of sample household data for the entire survey area. The independent variables generally considered are the income level, vehicle ownership, family size, type of dwelling unit, and trip length (Shuldiner, 1962; Fleet, Stowers and Swedloff, 1965; Chatterjee and Khasnabis, 1973). In some studies the distribution of households is related to certain zonal parameters and the theoretical possibility functions of the Gamma, Poisson and Binominal distributions are fitted with the observed data on the distribution of households by income, family size, and the number of employed residents respectively (Burns, Golob and Nicholaidis, 1976; Lerman and Moshe, 1976). Some models for estimating household characteristics of a study area are designed specially for

developing a three way classification of households based on household size, annual income, and vehicle ownership (Chatterjee, Khasnabis and Slade, 1977).

The advantages of these models are that they take full account of the variability of the trip data at household level and are independent of the zone system. The resulting equations are readily applicable to any system of areal units and the ecological problems are also reduced as the aggregation level has been brought down to the household level (Fleet and Robertson, 1968; McCarthy, 1969; Douglas and Lewis, 1971). Household-oriented disaggregate models were developed several years ago, but the task of forecasting household characteristics with these disaggregate models have not received much attention. This lack of interest may be attributed to the fact that neither the land use nor transportation planners have assumed the primary responsibility of this task.

2.1.4 Individual/Person Trip Models

These models are based on person trips and characteristics rather than household or zonal characteristics. Traditional approach of multiple regression analysis is generally used to explain trip productions. In a model (Chari, 1976) proposed for a medium sized Indian city, population is disaggregated into male and female groups. Further the male population is sub-divided into employed, self employed, school and college going persons. The large number of socio-economic variables are compressed into different basic variables to reflect the prosperity, stage in life cycle, mobility, opportunity, and

location. It is claimed that much of the information contained in these traits could be characterized by different proxy variables such as social status index, age, type of vehicle accessible and distance to home from Central Business District(CBD).

The basic approach taken in the household has been applied in an analogous way to a trip generation model with individual as the basic unit of analysis (Lansing, 1968; Tardiff, 1975 and 1977). The logit method for trip generation model has been developed by assuming an individual to make decisions on whether to make no trip, one trip etc., in a given period (Stopher and Lisco 1970; Talvitie, 1973).

Unlike the case of disaggregate mode choice models, disaggregate trip generation models have received little empirical or theoretical attention. Further the approach in most models is conceptually quite different from that taken in mode choice modelling. Whereas the latter is based on assumptions of individual behaviour, trip generation is usually thought to be an activity of the household rather than the individual (Oi and Shuldiner, 1962; Kassof and Deutchman, 1969; Smith and Cleaveland, 1976). The effect of different levels of spatial mobility is regarded as an attribute of individuals which arises as a product of both their individual circumstances (social and economic) and the accessibility of their home location to desired opportunities (Christopher Doubleday, 1977).

2.1.5 Scope of Trip Productions in the Study

To make reasonable forecasts of travel in subsequent years, it is essential to develop representative trip generation models from base year survey data. The increased urbanization of Indian cities has made the development of the trip generation models more desirable. To this date not much research has been done in this direction for Indian cities. It is proposed in this study to construct alternative trip production models for large to medium sized Indian cities, keeping in mind the availability of limited data. Of the four models described earlier (sub-sections 2.1.1 to 2.1.4) only three, namely, zonal regression models, category analysis, and household regression models are considered for construction. The individual/person trip models are not analysed because of difficulty of allocating household characteristics amongst the individuals. The three models are calibrated with the household survey data of Kanpur, a major industrial metropolis in the Northern India. The construction of zonal and household models require multiple regression technique. This technique is very powerful and has to be used in a judicious manner. A brief summary of multiple regression analysis is described in the Section 2.2.

2.2 MULTIPLE REGRESSION ANALYSIS

2.2.1 Assumptions

The principal assumptions of regression analysis (McCarthy , 1969; Neter and Wasserman, 1974; Hutchinson, 1974) are:

(i) the dependent variable is a linear function of the independent variables,

(ii) the independent variables are measured without error,

(iii) the deviation of the dependent variable about the regression line must be independent of each other and normally distributed,

(iv) the effects of the independent variables are additive and it is assumed that these are not correlated with one another, and

(v) the variance of the dependent variable about the regression line must be the same for all magnitudes of the independent variables.

It is difficult to ensure that the above basic assumptions are satisfied in most of the trip generation studies. Firstly the so called independent variables in the regression equations are not fully independent of each other and some sort of correlation exists among them. Secondly many of these variables, strictly speaking, are not normally distributed. Further some of them are not continuous. Inspite of all the drawbacks the multiple linear regression technique gained popularity because of its easy adoptability to computer programming when handling voluminous data.

2.2.2 Model Formulation

The typical multiple linear regression is of the form

$$Y = a_o + \sum_{i=1}^{n} a_i X_i + e \qquad .. (2.1)$$

where, Y is the variable under study and is known as the response or dependent variable and $X_1, X_2, \ldots, X_n$ are the variables which exhibit a causal effect on the value of Y and are known as the explanatory or independent variables. $a_1, a_2, \ldots, a_n$ are the least squares estimators of the unknown parameters and are known as the regression coefficients. a_o is a constant included to represent that portion of the value of Y not explained by the independent variables. ϵ is the error term.

In a typical regression analysis the given data relates to the present day values of the dependent variable Y and the independent variables, $X_1, X_2, \ldots, X_n$. The statistical technique of least squares fitting process is then applied to determine those values of the regression coefficients $a_1, a_2, .. a_n$ and the constant a_o which best fit the given data (Nie et al. 1977). The statistical validity of the model derived through multiple linear regression can be tested by the analysis of variance approach. The multiple correlation coefficient R indicates the degree of association between the dependent variable Y and the independent variables $X_1, X_2, \ldots, X_n$. It varies between 0 and 1.0; the closer it is to 1.0 the better the linear relationship. The closer R is to 0 the worse is the linear relationship (Bruton, 1975). The significance of R is that its square, the coefficient of determination, is approximately the decimal fraction of the variation in the dependent variable Y which is accounted for by the independent variables $X_1, X_2, .., X_n$. The standard error of estimate (SEE) indicates the degree of

variation of the data about the regression line established and is used to assess the value of the regression equation for prediction purposes. This statistic is sometimes referred to as the root mean square error or the residual standard deviation. It compares the accuracy of the value of Y' predicted for the present day situation using the regression equation with the observed values of Y which are used to derive the regression equation and is calculated as follows:

$$SEE = \sqrt{\frac{\Sigma(Y-Y')^2}{N}} \qquad (2.2)$$

where, SEE = standard error of estimate ,

Y = observed data used to derive regression equation,

Y' = value of Y calculated from the regression equation,

N = total number of observations.

The standard error of estimate (SEE) is more meaningful when it is expressed as a percentage of the mean observed value of the dependent variable Y (SEEP), i.e.,

$$SEEP = \frac{\text{Numerical SEE}}{\text{Mean observed value of the dependent variable}} \times 100 \qquad (2.3)$$

A good equation has a standard error of estimate which is a small percentage of the mean observed dependent variable and vice versa.

The regression sum of squares can be used to give some indication concerning whether or not the model is an adequate explanation of the true situation. The overall F-test uses statistical inference procedures to test the null hypothesis

that the multiple correlation coefficient is zero in the population from which the sample is drawn (Nie et al. 1977).

$$F = \frac{SSR/k}{SSE/(N-k-1)} \text{ at } k \text{ and } (N-k-1) \text{ degrees of freedom} \quad .. \quad (2.4)$$

where, SSR = regression sum of squares = $\Sigma(Y'-\bar{Y})^2$,

SSE = residual sum of squares = $\Sigma(Y-Y')^2$,

$\bar{Y}$ = mean value of Y ,

k = no. of independent variables entered in the equation,

N = total number of observations.

If the value of calculated F at k and (N-k-1) degrees of freedom is more than the critical value of F at k and (N-k-1) degrees of freedom even at 0.001 percent significance level, it may be concluded that it is unlikely that the sample is drawn from a population in which multiple R = 0.

The significance of the regression coefficient of each independent variable in a regression equation $a_1, a_2, ..., a_n$ is indicated by the t - test statistic. The value of t is calculated by dividing the regression coefficient by its standard error and must have a value of at least 2.0 for the significance to be established. Independent variables which have a value of t less than 2.0 do not have a significant relationship with the dependent variable and do not contribute any thing to the equation. The F-statistic can also be used as an alternative measure of the significance of the regression coefficient of each independent variable and is the square of the t-value, with degree of freedom 1 and(N-k-1) . All F-ratios should be compared to the

standard F-distribution with degrees of freedom of 1 and (N-k-1) at certain level of confidence (Neter and Wasserman, 1974; Bruton, 1975; Nie et al. 1977).

2.2.3 Selection of Best Set of Independent Variables

Selection of the best set of independent variables is highly essential in multiple regression analysis. A good search procedure must be able to handle important inter-correlated independent variables in such a way that not all of them will be removed. The two methods are:

(a) all possible regressions

(b) forward stepwise regression

All possible regressions : This procedure calls for the examination of all possible regression equations involving the potential independent variables and the selection of the best equation. The procedure is good when the number of independent variables are less. If the number of independent variables are more say p, the number of possible regressions are 2^p. This is not feasible if the number of independent variables are more in the model.

Forward stepwise regression: This procedure is developed to economize the computational work as compared with all possible regression approach while arriving at a reasonably good best set of independent variables. In this method it computes a sequence of regression equations at each step adding an independent variable. The criteria for adding an independent variable is decided on coefficient of partial correlation of F - value.

First the forward stepwise regression method calculates all simple regressions for each of the independent variables. For each of the simple equations, the F-statistic for testing whether the slope is zero is obtained by using the equation

$$F_k = \frac{MSR(X_k)}{MSE(X_k)} = \frac{\text{Mean square due to regression for independent variable } k}{\text{Mean square due to residuals}} \quad .. \quad (2.5)$$

The independent variable with the largest F-value/max correlation coefficient is the first to enter. If this value exceeds a predetermined level, the independent variable is entered. Otherwise the process terminates with no independent variable considered sufficiently helpful to enter the regression model.

If X_1 is the independent variable entered at step 1, the stepwise regression model now calculates all regressions with the two independent variables, say X_1, X_k as one of the pair. For each such regression, the F-statistic is calculated by using the formula,

$$F_k = \frac{MSR(X_k/X_1)}{MSE(X_1, X_k)} = \left(\frac{\text{Coefficient of } X_k}{\text{Standard error of } X_k}\right)^2 .. (2.6)$$

The independent variable with the largest F is the candidate for addition at the second stage. If this F-value exceeds a predetermined level the second independent variable is added. Otherwise the process terminates. In this way the entire procedure will be repeated and the process terminates when the F-value for all the variables in the model is above the predetermined level.

There are three methods for comparing the fitted models. They are:

(i) coefficient of multiple determination(R_p^2)criterion,

(ii) mean square error (MSE_p) criterion,

(iii) total square error (C_p) criterion.

The subscript p indicates that there are p parameters or (p-1) independent variables.

Coefficient of multiple determination (R_p^2) criterion: In this, the coefficient of multiple determination is compared to select the best set of independent variables. The drawback of this method is that this value does not take into account, the number of variables in the model since this value goes on increasing as the number of independent variables increases.

Mean square error (MSE_p) criterion: This method considers the number of parameters in the model into account through the degrees of freedom. Indeed minimum (MSE_p) can increase as the number of parameters p increases, if the reduction in residual sum of squares (SSE_p) becomes so small that it is not useful to add some more independent variables. When this method is used the set of independent variables which minimizes mean square error or a set for which this value is so close to the minimum is selected.

Total square error (C_p) criterion: This method is concerned with the total squared error of the n fitted observations for any given regression model. The total square error is a combination of bias component and a random error component.

If there is no bias in a regression equation with (p-1) independent variables, C_p has an expected value of p. When C_p values for all possible regressions are plotted against p, those regressions with little bias will tend to fall near $C_p = p$. When this method is used, the set of independent variables that leads to the smallest C_p value should be selected.

2.3 DATA COLLECTION FOR CASE STUDY

2.3.1 Background

The system of models for transport planning process as formulated in Section 2.4 and in Chapters 3,4 and 5 need to be calibrated and validated with the real world data. As such the city of Kanpur has been selected for this study. Kanpur, the major industrial metropolis of the Northern India in the State of Uttar Pradesh, is situated on the bank of the river Ganges. Like other important urban centres, it has acquired great significance due to increase in population and also of tremendous growth of various commercial/industrial activities. Urban population has been growing at a rapid rate especially in developing countries (World Bank, 1975). Kanpur city has a population of around 1.7 million (Census, 1981) registering a phenomenal increase of 32.4 percent in the preceeding decade. The population growth along with rapid industrialization exerts enormous pressure on the comparatively ill-developed infrastructural facilities and more so in the transportation planning.

The surface transportation routes in Kanpur city account for about 12 percent of the developed area (World Bank, 1981) which is very less when compared with other metropolitan areas in the country. The city road system can be approximated to a spider web pattern. The road network for the city is as shown in Fig. 2.1. Inspite of being a major metropolis of Northern India, the traffic on the roads is a mixture of several modes ranging from very slow moving hand driven cart to the fast moving vehicles like cars, buses etc. The roads especially in the core area are quite narrow. The core area is very densely populated and even the godowns of trade and commerce establishments are located in the heart of the city. The middle and outer areas are reasonably well served.

The public transportation system mainly consist of cycle rikshaws, buses and tempos. The buses and tempos have very limited use due to narrow and crowded roads in the core area. They mostly ply on the major roads and connect the outer areas to the Central Business District (CBD). Cycle rikshaw is the main mode of public transport inspite of the fact that they are the slow moving vehicles. The bus transit system of the city is handicapped by many problems such as mixed traffic flow, narrow streets, high peddestrian impedence, very high incidence of bicycle and cycle rikshaw traffic. The heavy influx of slow moving vehicles on the streets has reduced the operating speed of automobiles to one fourth of their normal speed leading to congestion on the roads. The traffic volume

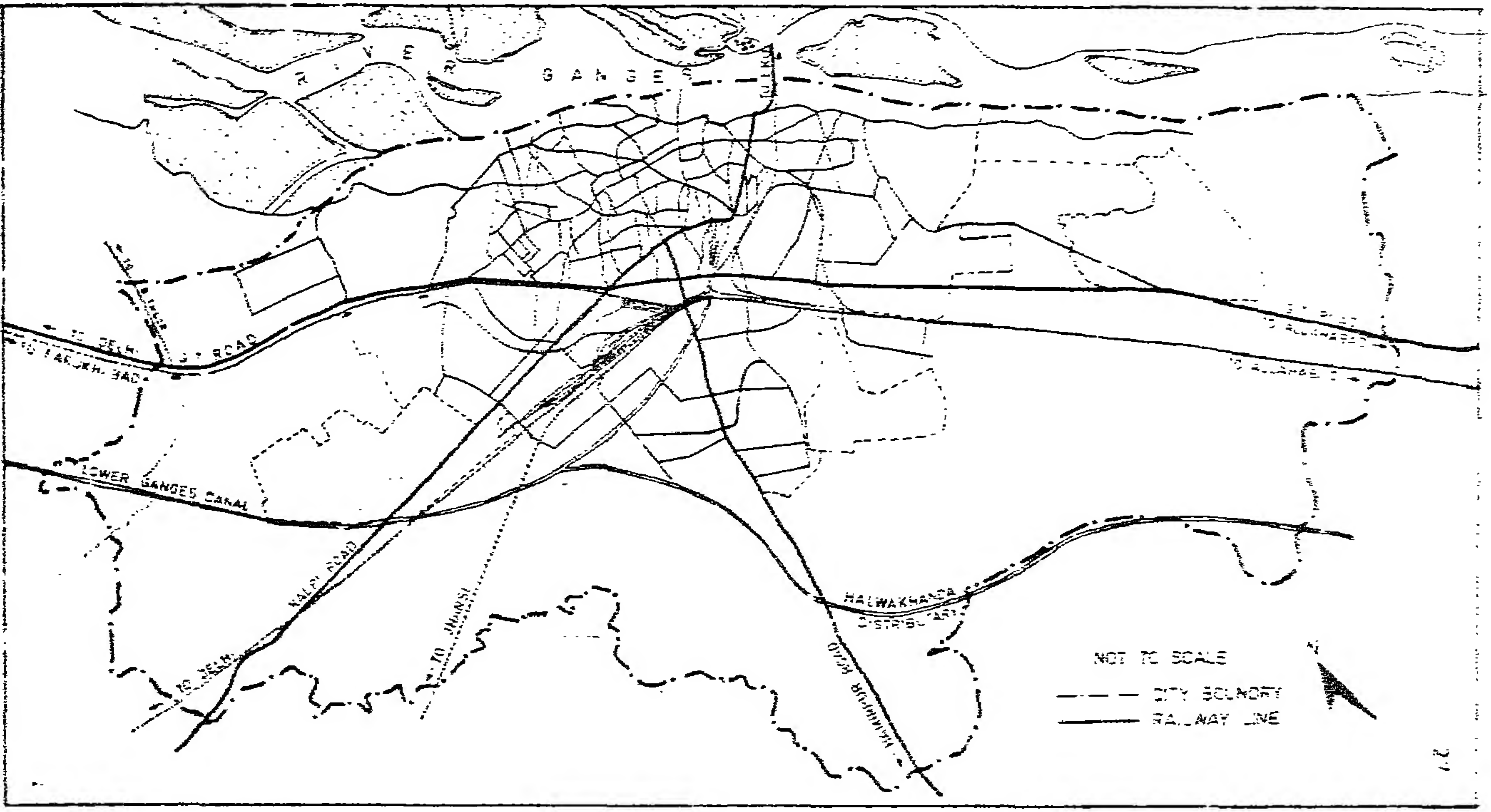


FIG. 2.1 KANPUR CITY ROAD NET WORK

on some of the roads in the core area is so high that the vehicles have to move almost in a state of congestion.

The population and other commerical and industrial activities of Kanpur are continuously increasing, whereas there has been practically no thinking about the planning of the transportation network. 'Lack of planning' concept can be said to be major cause of the sad state of the road traffic. The poor emphasis on transportation planning can be judged by the fact that there is notmuch information available about the existing conditions. Before constructing any model for the planning process it is necessary to conduct the studies related to the land use, socio-economic characteristics, the flow pattern of different categories of vehicles etc.

2.3.2 Household Survey

A household survey was carried out in 1978 to collect information about household characteristics and the existing travel pattern. The household survey was carried out by Town and Country Planning Department of the State of Uttar Pradesh. The study was however planned in association with the Indian Institute of Technology, Kanpur.

The city was divided into 3 cordons based on the land use activities and the population density. The innermost cordon, referred to as the central cordon in this study, describes the core area of the city. This core area is very densely populated with lot of trade and commerical establishments located in it. This in a way is the CBD of the city though

of a much larger size. This area has narrow streets in general and the flow of traffic on the major corridors is very heavy. The area between the central cordon and the middle cordon has low population density compared to the core area and includes small scale industrial units also. The economic level of the residents in this area is also better than those in the core area. The outer cordon defines the boundry of the metropolis. The area between the outer cordon and the middle cordon is thinly populated and the major industrial units, and large educational complexes are located in it. As there is hardly any open space available for new land use activities in and around core area, new residential locations are also coming up between middle and outer cordons.

As river Ganges is passing on the northern side of the city and it can be crossed only at two places, the development of the city is taking place in other directions only, especially on the southern side. The regions between the cordons were divided into various zones according to land use, density and economic standards. The whole city was thus divided into 56 traffic zones. The zones are homogeneous in character and are shown in Fig. 2.2 and names of the zones along with the cordon numbers are given in Appendix I. The streets form the boundry of a zone. In Kanpur the information related to number of households, population, and land use activity of an area is disaggregated to a level known as a 'CHAK' or 'WARD'. Each of the traffic zones contains a number of 'CHAKS'. While planning the area of a traffic zone it was seen that a 'CHAK'

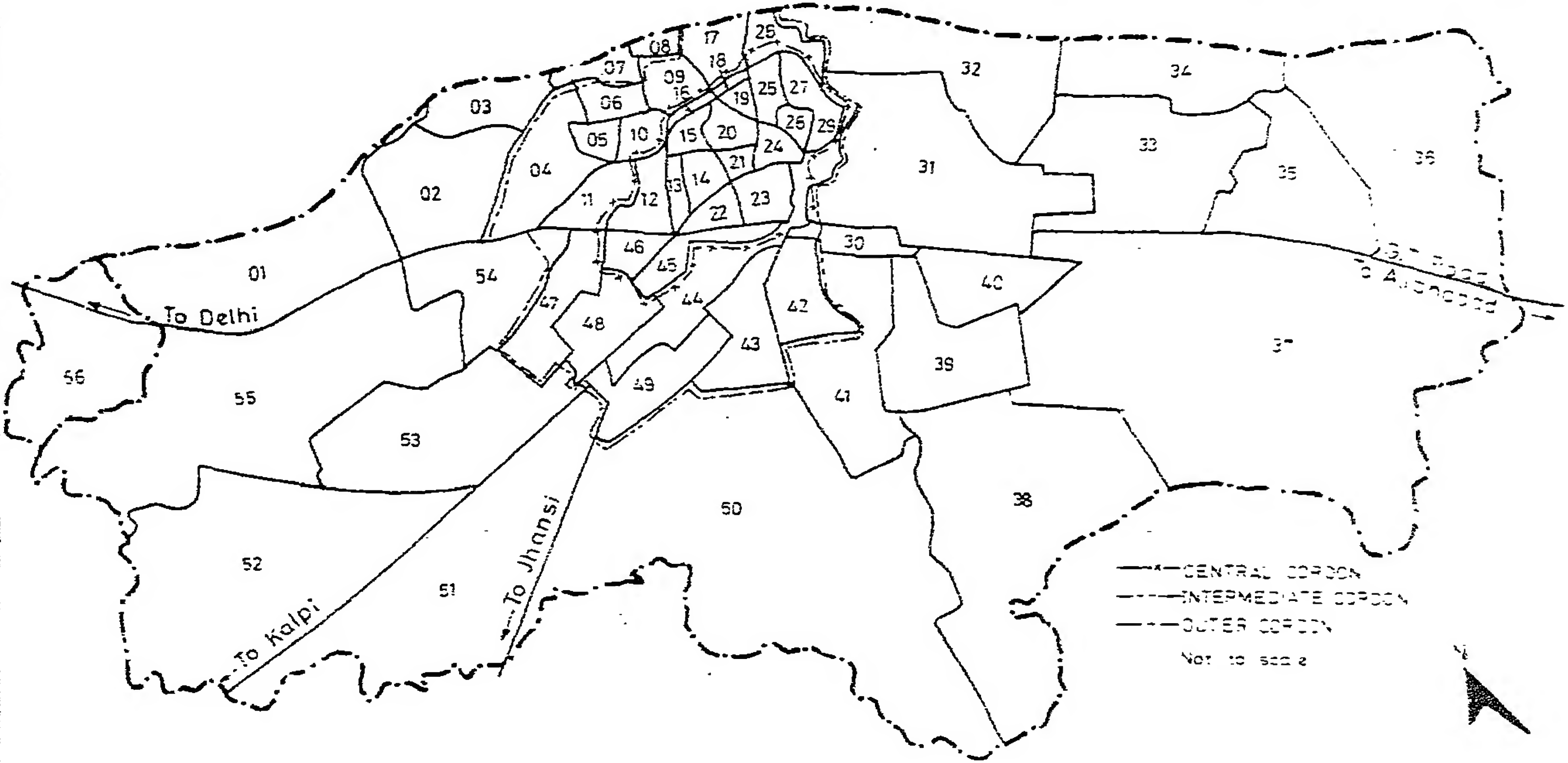


FIG-2-2 ZONING OF KANPUR CITY

is not divided between traffic zones.

The household survey was carried out only in 48 of the 56 traffic zones leaving some zones in the outer area due to various constraints. The survey was planned for a 5 percent sample size and in all 10471 households were surveyed. The choice of a household to be surveyed was made randomly from the household numbers available for each 'CHAK'.

The information gathered from the household survey related to the household characteristics and trip characteristics, is as follows:

Household characteristics:

- number of household members in each of the 5 age groups (0-4, 5-19, 20-24, 25-59, 60 and above);
- number of workers in the household;
- number of workers in the household, if any, going outside the city limits for the work;
- number of students in the household stratified into primary, secondry, degree, and vocational educational levels;
- number of vehicles of different types owned by the household;
- monthly household family income. This was stratified into 4 groups

  - lower income level (< Rs. 300)
  - lower middle income level (Rs.300 - Rs.1000)
  - upper middle income level (Rs.1000 - Rs.2000)
  - high income level (> Rs. 2000).

Trip characteristics

The following information related to the trips made on the last working day are recorded;

- purpose of trip;
- destination of the trip;
- trip length stratified into 5 groups;
    - upto 1 kilometer
    - 1-2 kilometer
    - 2-3 kilometers
    - 3-5 kilometers
    - more than 5 kilometers
- time of going and returning for home based trips;
- mode of travel
    - walk
    - bicycle
    - motorcycle/scooter
    - car
    - bus
    - cycle rickshaw.

The information relating to regular trips like work and education was well recorded. However the information related to other trip purposes like shopping, recreational and shopping etc. could not be fully recorded due to various constraints. The analysis of the household survey is thus restricted to only work and education trips. This survey of 10471 households produced information for a total of 21437 trips of which 12485 are the work trips, and the remaining 8952 are the education trips. The distribution of trips by different modes are given in Table 2.1.

TABLE 2.1 : DISTRIBUTION OF MODE CHOICE FOR SAMPLE HOUSEHOLD SURVEY DATA

Mode of travel	Number of work trips	Number of education trips	Total number of trips
Walk	5953	6844	12797
Bicycle	4967	812	5779
Motorcycle/ Scooter	762	31	793
Car	93	13	106
Bus	298	383	681
Cycle Rikshaw	412	869	1281

Table indicates that majority of the trips are made on foot only. Bicycle constitutes the next highly used mode of travel.

2.3.3 Travel Time Survey

Trip distribution, mode choice, and traffic assignment models require as input the travel time/travel cost by mode on various links of the network and also for various interzonal transfers. This is determined in this study by recording the travel time observations on different links. For each mode , number of travel time observations at peak periods are made and the most probable values are selected. The controids of the zones are identified and the interzonal travel times are also determined.

2.4 TRIP PRODUCTION MODELS FOR KANPUR CITY

2.4.1 Introduction

Development of trip production models involves a proper choice of independent variables. This choice is sometimes restricted to the variables for which data are available. Using the household survey data explained in Section 2.3, three trip production models, namely, zonal regression, category analysis, and household regression models are constructed for the city of Kanpur. The individual/person trip model is not analysed because of difficulty of allocating household characteristics amongst the individuals.

For selecting the best set of independent variables in the zonal and household regression models stepwise regression method is used. As the number of independent variables in the model increase, the coefficient of multiple determination R^2 also increases. The increase in R^2 is very significant at the initial stage and after sometime the increase in R^2 is very little, indicating that the inclusion of these independent variables which contribute very little to value of R^2 is not economical. Due to the difficulty in fitting the model based on R^2 value, it is decided to compare the fitted model by making use of MSE criteria (sub-section 2.2.3) and consider the model with the independent variables only upto the point where MSE/SSE is minimum.

2.4.2 Zonal Regression Models

These models explain the zonal trips in terms of zonal characteristics. Zonal values for the dependent travel variables are obtained by aggregating trip information from the household surveys into trip totals for each zone. The trips are specified by purpose into two broad categories, namely, work trips and education trips. Trips by other purpose are not analysed because of limited information available from household survey. Trip production relationships are developed firstly for the total zonal trip productions. The zonal trips are disaggregated into six different modes of travel and zonal trip production relationships are also obtained for these. The independent variables, the choice of which depend on the type of data available, considered are:

(i) total number of persons in the zone (TNP) ;

(ii) total income level for the zone (TIC) -this is the weighted sum of the household income levels;

(iii) number of private cars owned in a zone (TC);

(iv) number of two wheelers (motorcycles/scooters) in a zone (TMC);

(v) number of bicycles in a zone (TBC);

(vi) total distance travelled by the trip makers of the zone (TTL) - this, like the total income level for the zone, is weighted sum of the trip lengths.

Two sets of zonal regression equations are developed. In one only the first five independent variables are included and in the second all the six variables are included. The dependent variables considered for the trip productions are:

(i) Work trips

- total productions with all modes combined (TWM),
- productions by each mode (WM_i, $i = 1,2,\ldots,6$)

(ii) Education trips

- total productions with all modes combined (TEM)
- productions by each mode (EM_i, $i = 1, 2,\ldots,6$)

Zonal trip productions:

Relationships for the zonal trip productions by work and education as obtained by stepwise regression method are given in Table 2.2. These relations indicate that 98 percent of the total variation is explained by regression for work trips in both equations. For education trips the percent of total variation explained is 94 to 96 percent. Such a high value of multiple correlation coefficient indicates the good validity of the models. The number of zonal work trips increase with the zonal population (TP), zonal income level (TIC), and the number of different types of vehicles owned in the zone. However the number of zonal work trips reduce with increase in the zonal trip length (TTL). This appears to be logical as in zones located away from CBD, the average trip length is generally more and increase in the zonal trip length may not significantly affect the zonal trip productions.

TABLE 2.2 : RELATIONS FOR ZONAL TRIP PRODUCTIONS

Trip purpose	Relation no.	Zonal regression equation	Multiple correlation coefficient (R)	Percent of total variation (R^2)
Work trips	1	$TWM = 0.30 + 0.20\,TNP + 0.19\,TIC + 1.41\,TC + 0.062\,TMC - 0.099\,TBC$	0.99	0.98
	2	$TWM = -37.27 + 0.20\,TNP + 0.22\,TIC + 1.63\,TC + 0.074\,TMC + 0.038\,TBC - 0.047\,TTL$	0.99	0.98
Education trips	1	$TEM = 76.17 + 0.064\,TNP - 0.21\,TIC + 2.05\,TC + 0.65\,TMC + 1.55\,TBC$	0.97	0.94
	2	$TEM = -41.60 + 0.11\,TNP - 0.34\,TIC + 1.12\,TC + 0.57\,TMC + 1.01\,TBC + 0.32\,TTL$	0.98	0.96

Each person increase produces 0.2 trips by both the relations. Addition of a car results in 1.41 to 1.63 zonal work trips. The value of the regression coefficients for other types of vehicles, namely, bicycles and motorcycles/scooters is quite low, but as the number of these vehicles owned in a zone is very high compared to the cars, they contribute significantly to the total work trip productions. Increase in household income levels also increases the work trips. This may be because in some households there are more than one worker and the entire family income level is recorded. The regression coefficients for the education trips indicate increase with zonal population (TNP) and number of vehicles of different types in a zone. However there is decrease in the education trips with the increase in zonal income level (TIC). For a typical industrial city of Kanpur this appears to be quite logical as households with low/lower middle income levels generally have large family sizes and hence more number of school going children.

The statistical validity of the trip production analyses is assessed by considering the standard statistical tests besides the multiple correlation coefficients. The standard error of estimate is used to assess the value of regression equation for prediction purposes. This estimate is more meaningful when it is expressed as a percentage of the mean observed value of the dependent variable (Bruton, 1975). Table 2.3 indicates that the standard error of estimate is only 10 percent of the mean work trips, whereas, it is 16 to 19 percent of the mean education trips. The size of the regression constant in

TABLE 2.3 : STATISTICAL ANALYSIS FOR ZONAL TRIP PRODUCTIONS

Trip purpose	Dependent variable	Relation no.	Mean dependent variable	Regression constant (a_0)	Standard error of estimate (SEE)	SEE as a percentage of mean dependent variable (SEEP)	Overall F-value	Critical F-value at 0.001 percent
Work trips	THM	1	5202	0.30	531.14	10.21	528.21	5.53
		2	5202	-37.27	527.49	10.14	446.43	5.12
Education trips	TEM	1	3730	76.17	740.02	19.84	148.20	5.13
		2	3730	-41.60	624.78	16.75	176.29	5.12

relation to the mean observed value of the dependent variable is quite low indicating the reasonableness of the regression relations.

The F-ratio statistic is used as an overall measure of significance of the regression coefficients and it is the square of the t-value with the degrees of freedom k and (N-k-1). The F-ratios obtained from regression relationships and the values of the F-statistic at k and (N-k-1) degrees of freedom at 0.001 percent significance level are given in Table 2.3. Reference to the F-distribution indicates that the probability of getting an F-ratio greater than the value obtained is less than 0.001. It may be concluded that it is unlikely that the sample is drawn from a population in which multiple correlation coefficient R, is equal to zero. These tests indicate the statistical validity of the trip production relations.

Zonal productions by individual modes:

The regression relations for work trip productions by different modes as obtained by stepwise regression procedure are given in Table 2.4. Two sets of relations, one excluding zonal trip length (TTL) as independent variable, and the other including it are obtained. The percent of total variations explained by regression is quite high, generally between 83 to 93 percent, though in one case it is only 64 percent. It is further noted that the multiple correlation coefficient is higher when zonal trip length (TTL) is also included as an independent variable. It is only in one case of dependent variable (WM_6-work trips by cycle rikshaw), that the independent

TABLE 2.4 : RELATIONS FOR ZONAL WORK TRIP PRODUCTIONS BY INDIVIDUAL MODES

Type of mode	Relation no.	Zonal regression equation	Multiple correlation coefficient (R)	Percent of total variation (R^2)
(1)	(2)	(3)	(4)	(5)
Walk	1	$WM_1 = 211.38 + 0.27\,TNP - 0.022\,TIC + 0.93\,TC - 0.61\,TMC - 0.978\,TBC$	0.938	0.880
	2	$WM_1 = 73.65 + 0.25\,TNP + 0.085\,TIC + 1.73\,TC - 0.57\,TMC - 0.476\,TBC - 0.17\,TTL$	0.950	0.900
Bicycle	1	$WM_2 = -112.86 - 0.063\,TNP + 0.14\,TIC - 0.54\,TC - 0.29\,TMC + 0.95\,TBC$	0.986	0.970
	2	$WM_2 = -49.28 - 0.057\,TNP + 0.090\,TIC - 0.91\,TC - 0.31\,TMC + 0.72\,TBC + 0.079\,TTL$	0.990	0.980
Motor-cycle/ scooter	1	$WM_3 = -10.75 - 0.0026\,TNP + 0.014\,TIC - 0.078\,TC + 0.80\,TMC - 0.016\,TBC$	0.969	0.940
	2	$WM_3 = 2.11 + 0.14\,TC + 0.80\,TMC - 0.06\,TBC + 0.016\,TTL$	0.976	0.950

Contd...

variable of zonal income level (TIC) and the zonal population (TNP) do not enter the regression equation whereas all independent variables get included in the other dependent variables. This signifies that the choice of the independent variables is quite appropriate. Further the zonal work trips produced by bus (WM_5) decrease with increase in number of cars in the zone. This is quite logical as increase in number of cars distracts people from using public transport. Another significant observation is that increase in number of motor cars make a positive contribution to the work trips by walk. This may be because for short trip lengths people are attracted to walk to place of work inspite of the availability of the motor car due to high cost of fuel. The number of cars and motorcycles/scooters also make a positive contribution to the work trips by cycle rikshaw. This strengthens the above arguement that people owning cars are attracted to other public modes of transport due to high cost of operation of private vehicles and further due to congested flows on the city roads. The number of bicycles make a negative contribution to the cycle rikshaw trips, as any one owning a bicycle will not be attracted to cycle rikshaw as the two modes are having the same level of service.

The regression equations for education trips produced by different modes of travel are given in Table 2.5. The variation explained by regression is between 80 to 92 percent for the education trips performed on foot, bicycle, and cycle rikshaw. These three modes contribute a major portion of the

TABLE 2.5 : RELATIONS FOR ZONAL EDUCATION TRIP PRODUCTIONS BY INDIVIDUAL MODES

Type of mode	Relation no.	Zonal regression equation	Multiple correlation coefficient (R)	Percent of total variation (R^2)
(1)	(2)	(3)	(4)	(5)
Walk	1	$EM_1 = 102.2 + 0.0899\,TNP - 0.24\,TIC + 1.31\,TC - 0.54\,TMC + 1.25\,TBC$	0.957	0.920
	2	$EM_1 = 54.06 + 0.11\,TNP - 0.29\,TIC + 0.93\,TC - 0.057\,TMC + 1.03\,TBC + 0.13\,TTL$	0.957	0.920
Bicycle	1	$EM_2 = 1.42 - 0.0096\,TNP + 0.010\,TIC + 0.21\,TC + 0.27\,TMC + 0.13\,TBC$	0.869	0.750
	2	$EM_2 = -18.02 - 0.0017\,TNP - 0.0097\,TIC + 0.058\,TC + 0.25\,TMC + 0.044\,TBC + 0.053\,TTL$	0.896	0.800
Motor-cycle/ scooter	1	$EM_3 = 0.027 - 0.0011\,TNP - 0.0010\,TIC + 0.126\,TC + 0.011\,TMC + 0.012\,TBC$	0.690	0.480
	2	$EM_3 = -1.31 - 0.0006\,TNP - 0.0024\,TIC + 0.12\,TC + 0.010\,TMC + 0.0061\,TBC + 0.0036\,TTL$	0.728	0.530

Contd.....

Table 2.5 contd....

(1)	(2)	(3)	(4)	(5)
Car	1	$EM_4 = 0.369 - 0.00095\,TNP + 0.0025\,TIC + 0.16\,TC - 0.0022\,EMC - 0.0007\,TBC$	0.795	0.632
	2	$EM_4 = 0.063 - 0.00083\,TNP + 0.002\,TIC + 0.17\,TC - 0.0024\,EMC - 0.0020\,TBC + 0.00083\,TTL$	0.798	0.640
Bus	1	$EM_5 = -36.62 - 0.0059\,TNP + 0.0083\,TIC + 0.37\,EMC + 0.047\,TBC$	0.757	0.570
	2	$EM_5 = -53.74 - 0.0066\,TIC - 0.14\,TC + 0.35\,EMC - 0.030\,TBC + 0.47\,TTL$	0.800	0.640
Cycle rikshaw	1	$EM_6 = 8.71 - 0.0084\,TNP + 0.0093\,TIC + 0.24\,TC + 0.55\,EMC + 0.10\,TBC$	0.878	0.770
	2	$EM_6 = -22.61 + 0.0042\,TNP - 0.023\,TIC + 0.53\,EMC - 0.039\,TBC + 0.084\,TTL$	0.930	0.870

total education trips. For the education trips performed by car, bus and two wheeler (motorcycles/scooters) the variation explained by regression is between 50 to 60 percent which appears to be reasonable considering the small sample size of the dependent variable. It is observed that the independent variable, namely, number of private cars owned in a zone (TC) is not included in the education trips performed by bus and cycle rikshaw. This is quite logical as even in work trips there is no positive contribution by this variable. The zonal population (TNP) makes positive contribution in case of education trips performed on foot, whereas, it makes no significant contribution to trips by other modes. Increase in population of low/lower middle income zones generates education trips mostly of school going children. As the primary and secondry schools are generally located in close vicinity of the large residential locations, these trips are generally performed by children just by walk. The zonal income level makes a negative contribution to the education trips by walk, whereas, it behaves otherwise for other modes. This is because with increased income level children get attracted to other modes. It may also be noted that the inclusion of zonal trip length (TTL) increases the validity level of the equations.

When the mean observed value of the dependent variable is large, the standard error of estimate for a regression equation is only a small percent of the mean observed value. This is clear from Table 2.6. The standard error of estimate is a large percent of the observed value when the observed value of the

TABLE 2.6 : STATISTICAL ANALYSIS OF ZONAL TRIP PRODUCTIONS BY INDIVIDUAL MODES

Trip purpose	Type of mode	Dependent variable	Relation no.	Mean dependent variable	Regression constant (a_0)	Standard error of estimate (SEE)	SEE as a percentage of mean dependent variable (SEEP)	Overall F-value	Critical F-value at 0.001 percent
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Work trips	Walk	WM_1	1	2481	211.38	734.35	29.61	62.10	5.53
			2		73.65	647.22	26.09	68.90	5.12
	Bicycle	WM_2	1	2069	-112.86	317.42	15.35	311.50	5.53
			2		- 49.28	265.07	12.87	375.40	5.12
	Motor-cycle/scooter	WM_3	1	318	- 10.75	77.39	24.35	130.20	5.53
			2		2.11	67.64	21.28	216.00	6.12
	Car	WM_4	1	39	0.57	22.68	58.80	92.66	5.53
			2		- 0.90	22.48	58.27	78.86	5.12
	Bus	WM_5	1	124	- 58.43	110.14	88.62	16.30	5.53
			2		- 28.22	77.68	62.56	34.40	5.12
	Cycle rikshaw	WM_6	1	171	- 29.60	120.54	70.32	14.90	5.53
			2		- 34.46	121.31	70.77	12.30	5.12

Contd....

TABLE 2.6 contd...

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Education trips	Walk	EM_1	1	2851	102.20	723.89	25.39	92.66	5.53
			2		54.06	713.42	25.02	79.91	5.12
	Bicycle	EM_2	1	339	1.42	163.05	48.26	25.91	5.53
			2		-18.02	148.55	43.83	27.83	5.12
	Motor-cycle/scooter	EM_3	1	13	0.027	15.12	-	7.66	5.53
			2		- 1.31	14.50	-	7.72	5.12
	Car	EM_4	1	6	0.369	10.79	-	14.47	5.53
			2		0.063	9.29	-	12.04	5.12
	Bus	EM_5	1	159	-36.62	147.12	92.37	14.44	5.53
			2		-53.74	135.70	85.22	15.28	5.53
	Cycle rikshaw	EM_6	1	362	8.71	165.75	45.77	28.31	5.53
			2		-22.61	126.64	34.97	54.46	5.53

dependent variable is rather very small. F-ratio statistics given in Table 2.6 show that the probability of getting an F-ratio greater than the value obtained is less than 0.001. It is thus unlikely that the sample is drawn from a population in which multiple correlation coefficient R is equal to zero.

Discussion

The regression equations constructed for home based work and education trips are statistically quite reliable as is evident from the tests. As these types of relations have not been attempted before for large and medium sized cities of India, an assessment of temporal stability could not be attempted. When O-D survey data are available for later time periods it will be possible to conduct tests of stability.

Zonal regression is simple, as the travel data is aggregated and made quite manageable. Aggregation of the data makes it smoother than the original one. This may exhibit a misleading high level of correlation between the variables. Aggregation of data also implies homogeneity within the zones. Zonal equations do not provide an explanation of trip production at levels other than the zonal aggregated data on which they are based.

2.4.3 Category Analysis

Lack of homogeneity within the zones has encouraged researchers to adopt original or disaggregated data at household/ person level for trip generation. When household is taken to be

the behavioural unit, sample households are classified by type to conserve homogeneity within types. Each type of household is characterized by its mean trip rate. This technique known as category or cross classification analysis is also being used for trip attractions.

An attempt has been made in this study to perform category analyses of the work and education trip productions for the city of Kanpur. Considering the type of household data available and with the aim to conserve homogeneity within each type of household, a household is classified in three ways into the following subgroups.

(i) Family size classification

The size of the family significantly affects the trip productions. The household is divided into the following four subgroups:

- family upto 4 members
- family with 5 members
- family with 6 members
- family with more than 6 members

(ii) Income level classification

Four subgroups of this type are:

- low income group
- lower middle income group
- upper middle income group
- high income group

Household survey has considered only these four groups and they are adopted as such.

(iii) Vehicle ownership classification

The number and type of vehicle in the household affects the trips produced by different modes and also to some extent the total trips produced. Preliminary analysis of the sample data is done to investigate the distributions of the vehicles owned. Based on that analysis the households have been classified into the following eight subgroups of vehicle ownerships:

- no vehicle
- bicycle only
- motorcycle/scooter only
- car only
- bicycle and motorcycle/scooter
- bicycle and car
- motorcycle/scooter and car
- bicycle, motorcycle/scooter and car.

The cross classification of the households have been done in three different ways, each in terms of the following two variables:

(a) family size - income level

(b) family size - vehicle ownership

(c) income level- vehicle ownership

These three classifications have been chosen to study the variations in each stratification. All the households interviewed in the travel survey are sorted into one of the cells in each classification. The number of trips produced by the households in a cell and also the number of households in a cell are summed up and the trip rate for each cell is established. The category analysis of the total trip productions (including work,

and education trips) by each of the above three mentioned sets are given in Tables 2.7 to 2.9.

Table 2.7 and Fig. 2.3 of family size - income level classification indicate that the household trip production rate increases with the size of the family. This is quite logical because a large family produces more work and education trips. Further in a particular family size the trip production rate also increases with income level. This may be because the total household income is recorded in survey. When a household has more than one worker, the recorded income level goes up and the work trip productions also increase . In case of individuals having high income level, the work and education production rate are generally also higher. Table 2.8 and Fig.2.3 of family size - vehicle ownership again establish that increased number of persons in a household increases the trip production rate. Further the number and type of vehicles owned by the household also contribute to trip productions. Households having no personal vehicle produce minimum trip rate. For the households having one personal vehicle, the type of vehicle does not significantly affect the production rate. Household with two or more vehicles produce highest trip rate. This category analysis closely resembles with the trend in the previous category of family size - income level because the income level and type of vehicles owned are highly correlated.

Trip production rates for income level vehicle ownership category are also analysed in Table 2.9 and Fig. 2.4. In this category some cells of the matrix have insignificant numbers.

TABLE 2.7 : HOUSEHOLD DAILY TRIP PRODUCTION RATES FOR FAMILY SIZE - INCOME LEVEL CATEGORY

Income level group	Family size (number of members)			
	upto 4	5	6	more than 6
Low income	1.187* (56260)**	1.953 (11580)	2.318 (7300)	2.760 (8140)
Lower middle income	1.570 (52880)	2.450 (21740)	2.830 (18480)	3.340 (25540)
Upper middle income	1.930 (2080)	2.644 (1180)	3.050 (780)	4.170 (2120)
High income	1.690 (520)	4.000 (160)	3.250 (160)	4.360 (500)

* Trip production rate for population

** No. of households in the population.

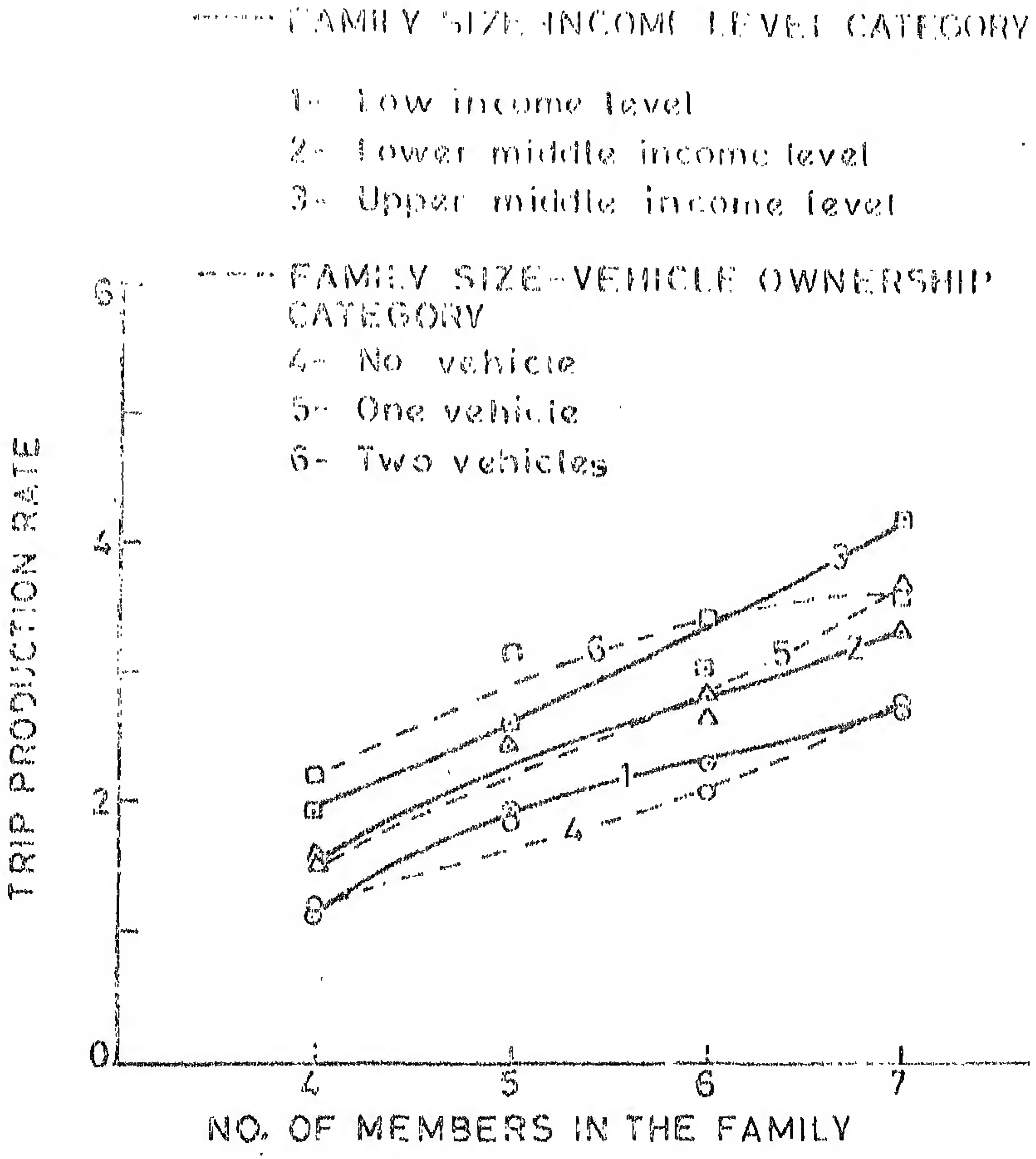


FIG.2·3 TRIP PRODUCTION RATES FOR FAMILY SIZE-INCOME LEVEL/VEHICLE OWNERSHIP CATEGORIES

TABLE 2.8 : HOUSEHOLD DAILY TRIP PRODUCTION RATES FOR FAMILY SIZE - VEHICLE OWNERSHIP CATEGORY

Vehicle ownership	Family size (number of members)			
	upto 4	5	6	more than 6
No vehicle	1.193* (54240)**	1.880 (9480)	2.120 (7000)	2.800 (9680)
<u>One vehicle</u>				
Bicycle	1.530 (50260)	2.410 (21220)	2.870 (16900)	3.400 (22340)
Motorcycle/ scooter	1.620 (4840)	2.320 (2080)	2.450 (1500)	3.110 (1880)
Car	1.474 (380)	2.600 (100)	-	4.500 (40)
<u>Two vehicles</u>				
Bicycle and motorcycle/ scooter	2.600 (100)	2.800 (100)	3.000 (80)	3.600 (100)
Bicycle and car	2.100 (1700)	3.040 (1480)	3.450 (1060)	4.170 (1800)
Motorcycle/ scooter and car	2.000 (200)	3.800 (100)	4.000 (60)	3.200 (200)
<u>Three vehicles</u>				
Bicycle, motor-cycle/scooter and car	-	3.600 (100)	3.200 (100)	4.380 (260)

* Trip production rate for population

** No. of households in the population

TABLE 2.9 : HOUSEHOLD DAILY TRIP PRODUCTION RATES FOR INCOME LEVEL - VEHICLE OWNERSHIP CATEGORY

Vehicle ownership	Income level group			
	Low income	Lower middle income	Upper middle income	High income
No vehicle	1.300* (44920)**	1.860 (34280)	2.540 (1000)	2.400 (200)
One vehicle				
Bicycle	1.830 (37020)	2.490 (70500)	3.260 (1560)	2.500 (240)
Motorcycle/ scooter	1.570 (920)	2.120 (7860)	2.640 (1380)	3.130 (140)
Car	-	1.750 (240)	2.080 (240)	2.330 (60)
Two vehicles				
Bicycle and motorcycle/ scooter	-	2.860 (140)	2.620 (160)	4.000 (80)
Bicycle and car	2.320 (380)	3.160 (4060)	3.540 (1360)	3.000 (240)
Motorcycle/ scooter and car	-	2.800 (100)	3.000 (280)	3.130 (160)
Three vehicles				
Bicycle, motor-cycle/scooter and car	-	4.330 (60)	2.890 (180)	4.450 (220)

* Trip production rate for population

** No. of households in the population.

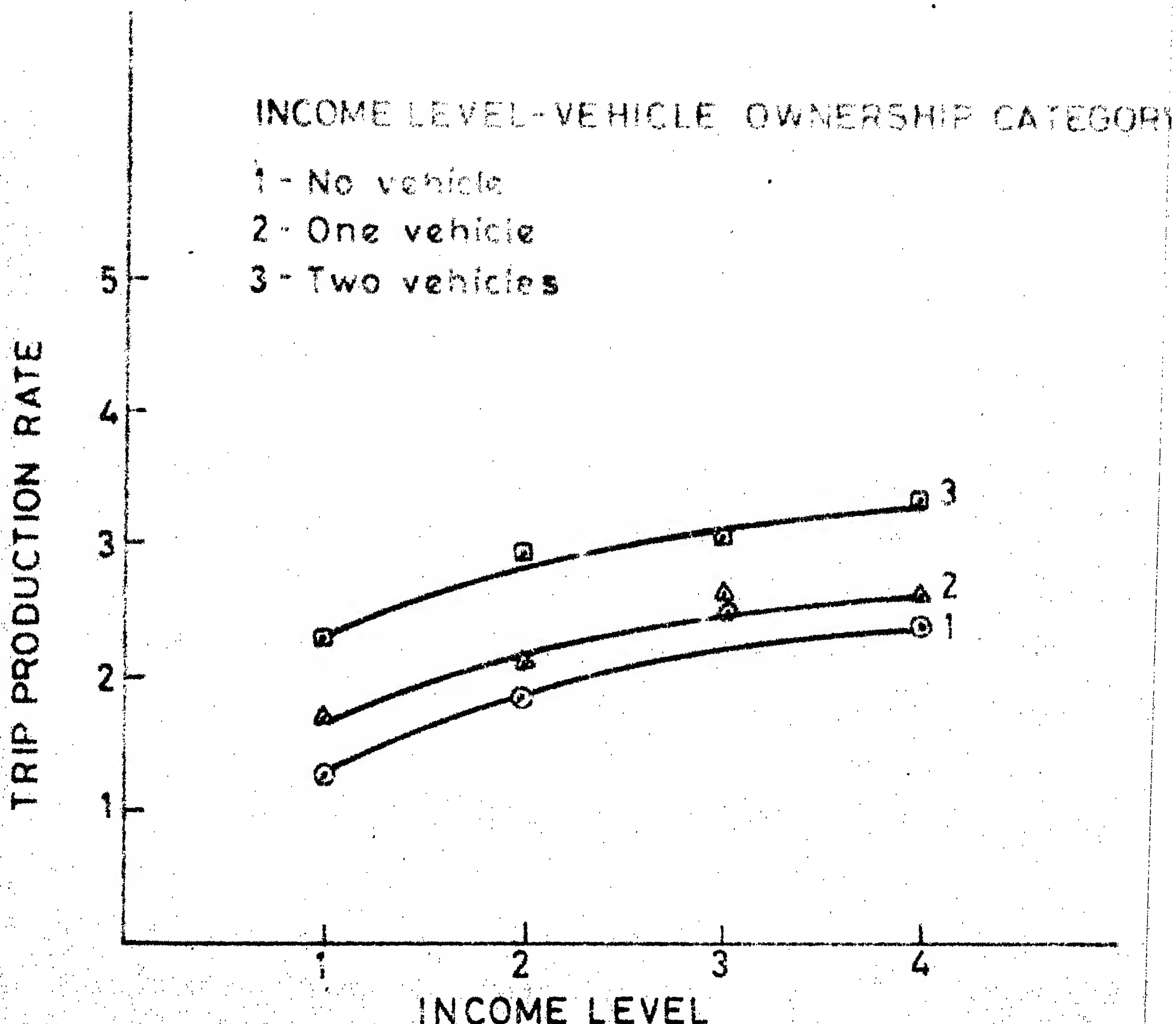


FIG.2·4 TRIP PRODUCTION RATES FOR INCOME LEVEL-VEHICLE OWNERSHIP CATEGORY

Further as mentioned earlier both the variables are correlated. The general trend of trip rates is similar to that explained previously for the two categories.

Category analyses of the trip production is also done seperately for work and education trips and the results are presented in Tables 2.10 to 2.12. Table 2.10 of the family size-income level classification for work and education trips shows that the household trip production rate increases with the income level in case of work trips. But the education trip production rate is highest for the lower middle income level. The low and lower middle income level people are having large family size in a city like Kanpur and hence the household trip rate is highest . The family size-vehicle ownership classification for work and education trips are shown in Table 2.11. Here the household trip rate increases with the vehicle ownership and is giving highest trip rate when the vehicle ownership is more. This is quite logical as the increase in vehicle ownership definitely increases the household trip rate. Table 2.12 of the incom level-vehicle ownership classification shows that household trip rate increases with the income level for both work and education trips. This may be because the households with high income level make more trips. When there are more number of workers in a household, the income level and hence the trip production rate increases. Category analysis is also done of the trips produced by different modes for each of the trip purpose. The trends are of similar nature as for the total trip production rate.

TABLE 2.10 : HOUSEHOLD DAILY TRIP PRODUCTION RATE BY PURPOSE FOR FAMILY SIZE - INCOME LEVEL CATEGORY

Income level group	Family size (number of members)							
	upto 4		5		6		more than 6	
	Work trips	Education trips	Work trips	Education trips	Work trips	Education trips	Work trips	Education trips
Low income	0.96* (56260)**	0.23 (56260)	0.98 (11580)	0.97 (11580)	1.02 (7300)	1.30 (7300)	1.11 (8140)	1.65 (8140)
Lower middle income	1.19 (52880)	0.38 (52880)	1.16 (21740)	1.29 (21740)	1.20 (18480)	1.63 (18480)	1.36 (25540)	1.98 (25540)
Upper middle income	1.66 (2080)	0.27 (2080)	1.45 (1180)	1.19 (1180)	1.46 (780)	1.58 (780)	2.05 (2120)	2.12 (2120)
High income	1.62 (520)	0.08 (520)	3.25 (160)	0.75 (160)	1.87 (160)	1.37 (160)	2.60 (500)	1.76 (500)

* Trip production rate for population

** No. of households in the population.

TABLE 2.11 : HOUSEHOLD DAILY TRIP PRODUCTION RATE BY PURPOSE FOR FAMILY SIZE - VEHICLE OWNERSHIP CATEGORY

Vehicle ownership	Family size - (number of members)							
	upto 4		5		6		more than 6	
	Work trips	Education trips	Work trips	Education trips	Work trips	Education trips	Work trips	Education trips
No vehicle	1.02* (54240)**	0.17 (54240)	1.01 (9480)	0.88 (9480)	1.12 (7000)	1.08 (7000)	1.20 (9680)	1.59 (9680)
One vehicle								
Bicycle	1.22 (50260)	0.41 (50260)	1.12 (21220)	1.29 (21220)	1.17 (16900)	1.69 (16900)	1.38 (22340)	2.01 (22340)
Motorcycle/ scooter	1.15 (4840)	0.47 (4840)	1.13 (2080)	1.19 (2080)	1.06 (1500)	1.39 (1500)	1.29 (1880)	1.82 (1880)
Car	1.26 (380)	0.21 (380)	1.20 (100)	1.40 (100)	-	-	1.50 (40)	3.00 (40)
Two vehicles								
Bicycle and motorcycle/ scooter	1.40 (100)	1.20 (100)	1.20 (100)	1.60 (100)	1.50 (80)	1.50 (80)	2.00 (100)	1.60 (100)
Bicycle and car	1.49 (1700)	0.61 (1700)	1.59 (1480)	1.44 (1480)	1.32 (1060)	2.13 (1060)	1.63 (1800)	2.53 (1800)
Motorcycle/ scooter and car	1.26 (200)	0.21 (200)	1.20 (100)	1.40 (100)	-	-	1.50 (200)	3.00 (200)
Three vehicles								
Bicycle, motorcycle/scooter and car	-	-	2.20 (100)	1.40 (100)	1.20 (100)	2.00 (100)	2.38 (260)	2.00 (260

* Trip production rate for population

** No. of households in the population.

TABLE 2.12: HOUSEHOLD DAILY TRIP PRODUCTION RATE BY PURPOSE FOR INCOME LEVEL – VEHICLE OWNERSHIP CATEGORY

Vehicle ownership	Income level group							
	Low income		Lower middle income		Upper middle income		High income	
	Work trips	Education trips	Work trips	Education trips	Work trips	Education trips	Work trips	Education trips
No vehicle	0.94* (44920)**	0.36 (44920)	1.16 (34280)	0.69 (34280)	1.94 (1000)	0.60 (1000)	2.00 (200)	0.40 (200)
One vehicle								
Bicycle	1.03 (37020)	0.79 (37020)	1.24 (70500)	1.25 (70500)	1.92 (1560)	1.33 (1560)	1.91 (240)	0.58 (240)
Motorcycle/ scooter	0.72 (920)	0.85 (920)	1.16 (7860)	0.96 (7860)	1.42 (1380)	1.22 (1380)	1.71 (140)	1.71 (140)
Car	-	-	1.33 (240)	0.42 (240)	1.17 (240)	0.92 (240)	1.33 (60)	1.00 (60)
Two vehicles								
Bicycle and motorcycle/ scooter	-	-	1.43 (140)	1.42 (140)	1.00 (160)	1.63 (160)	2.75 (80)	1.25 (80)
Bicycle and car	1.05 (380)	1.26 (380)	1.44 (4060)	1.72 (4060)	1.88 (1360)	1.66 (1360)	1.83 (240)	1.17 (240)
Motorcycle/ scooter and car	-	-	2.00 (100)	0.80 (100)	1.93 (280)	1.07 (280)	2.88 (160)	0.25 (160)
Three vehicles								
Bicycle, motor-cycle/scooter and car	-	-	1.33 (60)	3.00 (60)	1.22 (180)	1.66 (180)	3.00 (220)	1.45 (220)

* Trip production rate for population

** No. of households in the population

The trip production rates as estimated for each of the category are used to determine the zonal trip productions. These trip productions obtained from the models are compared with the observed zonal trip productions in Table 2.13. The estimated trip productions by the three models are quite close to the observed values. The variations are generally well within ± 10 percent and it is higher only when a zone produces very small number of trips. These results indicate that category analysis can be applied to estimate trip productions for a city like Kanpur.

Discussion

Cross classification technique provides an opportunity to examine the linkages between the purpose, mode of travel, household status and its mobility level. This technique, though very simple, leads to a better explanation of trip generation than obtained from aggregated data. One major limitation as observed from this study is the uneven cell size which makes some mean trip rates less reliable than others. It may be thus desirable to avoid large classification systems. In the absence of any other household survey data temporal stability of category analysis could not be attempted.

2.4.4 Household Regression Models

Household regression approach is a combination of the better features of zonal regression and category analysis. This approach applies regression analysis directly to the household characteristics instead of zonal characteristics as is done in zonal regression. It provides linear regression

TABLE 2.13 : COMPARISON OF OBSERVED AND MODEL ZONAL TRIP PRODUCTIONS

Zone No.	Observed trip productions	Model trip productions		
		Model I	Model II	Model III
(1)	(2)	(3)	(4)	(5)
1	1020	1180	1060	880
2	2900	3520	3280	3280
3	9140	9860	9960	9360
4	9620	8900	8800	9040
5	1900	1500	1500	1640
6	12340	11580	11680	11900
7	4040	4080	3860	3940
8	4940	4980	4740	4840
9	2720	2780	2720	2960
10	5900	5340	5540	5840
11	4800	4200	4200	4540
12	22760	21440	21920	22000
13	8320	7840	8400	7780
14	20460	19280	19720	19000
15	18840	19820	19820	18140
16	1100	940	920	940
17	5040	4860	4620	5000
18	360	440	440	460
19	5860	6680	6480	6160
20	14680	15140	15240	14460
21	8780	8980	9160	8300
22	3060	2940	3200	2820
23	7360	7260	7400	7040
24	11540	11440	11700	11720
25	17780	16820	16300	17040
26	1420	1220	1120	1280
27	17120	16880	17040	15780
28	9580	8200	7440	9000

contd....

TABLE 2.13 contd....

(1)	(2)	(3)	(4)	(5)
29	13740	13340	13780	13280
30	3980	3320	2980	3660
31	14760	17380	16820	16340
32	1920	2080	2032	2120
33	5660	6860	6720	6440
41	29160	32020	32550	31760
42	9680	8900	9080	9900
43	16500	16060	16240	16480
44	4240	4000	4200	4160
45	5640	5280	5380	5380
46	17320	16800	17000	17100
47	19240	17920	17780	19000
48	9240	8620	8340	9520
49	21020	20600	20400	20800
50	5120	6460	6300	5880
52	1640	2160	2000	2240
53	6560	7440	7400	8180
54	5080	5200	5380	4900
55	3420	4520	4420	4280
56	1440	1420	1400	1880

NOTE: MODEL I = FAMILY SIZE - INCOME LEVEL CATEGORY
MODEL II = FAMILY SIZE - VEHICLE OWNERSHIP CATEGORY
MODEL III= INCOME LEVEL- VEHICLE OWNERSHIP CATEGORY

equation instead of a catalogue of mean trip rates for different categories of households as is done in category analysis. The results of household regression are independent of the zone system and can be applied to any set of areal units. As the analysis is applied to the basic household data it operates on the whole of the variability which exists in the data at that level.

Household regression analysis is applied in this study for the city of Kanpur to the complete set of household data for the entire survey area. The choice of independent variables are guided by the results obtained earlier in zonal regression and household classification. The trip ends are stratified into total work and education trips and also by all vehicular modes as before. The explanatory variables that have been considered in the household regression method are:

(i) family size
(ii) income level
(iii) vehicle ownership
(iv) trip length.

These variables are further stratified into various subgroups as given below.

(i) Family size:

- family upto 4 members (FS_1)
- family with 5 members (FS_2)
- family with 6 members (FS_3)
- family with more than 6 members(FS_4).

(ii) Income level:

- low income group (IG_1)
- lower middle income group (IG_2)
- upper middle income group (IG_3)
- high income group (IG_4).

(iii) Vehicle ownership:

- no vehicle (V_1)
- bicycle (V_2)
- motorcycle/scooter (V_3)
- car (V_4)
- bicycle and motorcycle/scooter (V_5)
- bicycle and car (V_6)
- motorcycle/scooter and car (V_7)
- bicycle, motorcycle/scooter and car (V_8).

(iv) Trip length:

- upto 1 kilometer (TL_1)
- 1-2 kilometers (TL_2)
- 2-3 kilometers (TL_3)
- 3-5 kilometers (TL_4)
- more than 5 kilometers (TL_5).

In all there are 21 independent variables which are common to both work and education trips. However it may be possible that some independent variables have considerable influence, some others may explain the same effect and some may have only minor effect. Relationships are developed by making use of the stepwise regression method as explained in Section 2.2, by considering all the independent variables.

The regression equations of the different types of zonal trips produced as explained by the household characteristics of the zone are given in Table 2.14.

The equations indicate that not all the 21 explanatory variables are included due to the correlations among them. For the total work trip productions(Eqn.1 of Table 2.14)only 7 independent variables belonging to the trip length, income level and family size stratifications are included. No independent variablo of vehicle ownership type is included in this equation. However the regression equations for work trips by different modes do include the vehicle ownership variables. The maximum of 14 independent variables are included in the work trip production regression equations (Eqns.1 to 7) . Education trip regression equations (Eqns.8 to 14) generally include more number of independent variables,(9 to 15) than those for the work trips. Some independent variables do make negative contribution in certain cases which are generally found to be logical. The constant term in the equations may be regarded as an indication that other significant variables should have been included in the regression analysis.

The regression equation indicates a sufficiently high value of the multiple correlation coefficients both for work and education trips. These values are significantly higher than those obtained by zonal regression. The percent of total variation explained by regression is above 90 percent for work trips in all equations except one where it is 88 percent. The percent variation explained by regression for education trips

TABLE 2.14 : RELATIONS FOR HOUSEHOLD WORK AND EDUCATION TRIP PRODUCTIONS

Trip purpose and Eqn.no.	Dependent variable	Household regression equation	Multiple correlation coefficient (R)	Percent of total variation (R^2)
(1)	(2)	(3)	(4)	(5)
Work trips				
1	TWM (All modes)	$TWM = 7.34 + 0.95TL_1 + 0.97TL_3 + 0.96TL_4 + 0.83TL_5 + 0.086IG_2 + 0.046FS_1 + 0.078FS_4$	0.999	0.990
2	WM_1 (Walk)	$WM_1 = 70.76 + 0.11V_1 - 0.43V_2 - 0.73V_3 - 0.27V_4 - 0.91V_6 + 0.95TL_1 + 0.55TL_2 + 0.29TL_3 + 0.57TL_4 + 0.52TL_5 + 0.25IG_1 - 0.34FS_2$	0.995	0.990
3	WM_2 (bicycle)	$WM_2 = -23.76 + 0.80V_2 + 2.58V_5 + 0.35TL_3 + 0.10TL_4 + 0.17TL_5 - 0.29IG_1 - 0.62IG_3 - 0.059FS_1 + 0.19FS_4$	0.995	0.990
4	WM_3 (motor-cycle/scooter)	$WM_3 = -0.84 - 0.051V_2 - 0.95V_3 - 0.79V_4 + 2.07V_5 + 0.45V_6 + 1.23V_7 + 1.62V_8 - 0.23TL_2 - 0.006TL_5 + 0.056IG_2 - 0.37IG_3 - 0.25IG_4 + 0.15FS_2 - 0.067FS_4$	0.990	0.980
5	WM_4 (car)	$WM_4 = -0.15 - 0.029V_3 + 1.26V_4 + 0.44V_5 + 0.80V_7 + 1.15V_8 + 0.44TL_2 + 0.16TL_4 - 0.026TL_5 + 0.019FS_2 - 0.026FS_4$	0.970	0.940

Contd.......

TABLE 2,14 contd....

(1)	(2)	(3)	(4)	(5)
6	WM_5 (bus)	$WM_5 = -31.01 - 0.071V_2 + 0.81V_3 - 1.19V_5 - 0.032V_6 - 0.98V_7 - 0.92V_8 - 0.014TL_1 + 0.32TL_2 + 0.085TL_3 + 0.11TL_4 + 0.24TL_5 + 0.26IG_3 + 0.46IG_4$	0.936	0.880
7	WM_6 (cycle-rikshaw)	$WM_6 = -26.24 - 0.058V_2 - 1.48V_8 + 0.17TL_3 + 0.69TL_4 - 0.13TL_5 + 0.49IG_3 + 0.11IG_4 - 0.13FS_2 + 0.18FS_3$	0.900	0.810
Education trips				
8	TEM (all modes)	$TEM = 4.37 + 0.049V_2 - 0.31V_6 + 1.02V_8 + 1.04TL_1 + 1.17TL_2 + 1.01TL_3 + 1.08TL_4 - 0.059IG_1 + 0.11IG_3 - 0.30IG_4 - 0.16FS_2 + 0.10FS_3$	0.999	0.990
9	EM_1 (walk)	$EM_1 = 51.35 + 1.91V_1 - 1.16V_3 - 6.34V_5 - 1.47V_6 + 4.90V_7 - 2.17V_8 + 0.98TL_1 + 0.71TL_2 + 0.56TL_3 + 0.78TL_4 - 0.62TL_5 + 0.17IG_2 - 0.87FS_3 + 0.20FS_4$	0.999	0.990
10	EM_2 (bicycle)	$EM_2 = -7.69 + 0.0009V_2 + 1.47V_6 - 1.43V_7 + 0.11TL_3 - 0.17TL_4 + 1.03TL_5 + 0.077IG_1 - 0.17IG_3 - 0.102FS_4$	0.950	0.900

Contd.....

TABLE 2.14 contd...

(1)	(2)	(3)	(4)	(5)
11	EM_3 (motorcycle/ scooter)	$EM_3 = 0.79 + 0.35V_5 + 0.0077V_6 + 0.40V_7 + 0.0085TL_1 - 0.014TL_2 + 0.017TL_3 - 0.027TL_5 - 0.008IG_1 - 0.055IG_3 - 0.037FS_2 + 0.16FS_3 + 0.099FS_4$	0.870	0.760
12	EM_4 (car)	$EM_4 = -0.055 - 0.012V_3 + 0.14V_4 + 0.16V_5 + 0.22V_8 - 0.0014TL_1 + 0.14TL_4 - 0.025TL_5 + 0.037IG_3 - 0.0022FS_1 + 0.019FS_2 - 0.011FS_4$	0.870	0.760
13	EM_5 (bus)	$EM_5 = -27.69 + 0.55V_3 - 2.29V_4 + 4.46V_5 + 1.58V_8 - 0.050TL_1 + 0.11TL_3 - 0.061TL_4 + 0.44TL_5 + 0.16IG_1 + 0.12IG_2 - 0.13IG_3 - 0.093IG_4 - 0.19FS_1 + 0.26FS_3 - 0.20FS_4$	0.970	0.930
14	EM_6 (cycle rikshaw)	$EM_6 = -5.94 - 0.068V_1 + 0.66V_3 + 1.39V_4 - 0.88V_5 - 3.72V_7 + 0.026TL_1 + 0.23TL_3 + 0.33TL_4 + 0.30TL_5 - 0.074IG_2 + 0.68IG_3 + 0.0016FS_1 + 0.28FS_3$	0.970	0.940

NOTE: TWM : number of work trips produced in the zone

TEM : number of education trips produced in the zone

Contd...

TABLE 2.14 contd...

WM_i = number of work trips produced in the zone by ith mode ($i=1,2,\ldots,6$)

EM_i = number of education trips produced in the zone by ith mode ($i=1,2,\ldots,6$)

FS_k = number of households having family size of kth group ($k=1,2,\ldots,4$)

IG_l = number of households having income level of lth group ($l=1,2,\ldots,4$)

V_m = number of households having vehicle ownership of mth group ($m=1,2,\ldots,8$)

TL_n = number of households having trip length of nth group ($n=1,2,\ldots,5$).

is above 90 percent for five cases and it is only 76 percent for two cases.

Table 2.15 shows that the value of the regression constant in all the equations is quite small compared to the observed zonal mean of the dependent variables. This indicates the reasonableness of the regression analysis. The standard errors of estimate are also a small percent of the mean observed values. It is only in those cases where the mean observed value is quite low that the standard error of estimate is a little higher. F-ratio statistics of regression relationships and the value of the F-statistic at 0.001 percent significance level are also given in Table 2.15. Reference to the F-distribution shows that the probability of getting an F-ratio greater than the value obtained is less than 0.001. It may, therefore, be concluded that it is unlikely that the sample is drawn from a population in which multiple correlation coefficient (R) is equal to zero. These tests indicate the statistical validity of the household trip production relations.

Discussion

The household regression analysis has produced sufficiently encouraging results to be preferred over zonal regression and household classification. This approach takes full account of the variability of the trip data at household level and provides a better understanding of trip productions than gained from mean trip rates corresponding to category analysis.

TABLE 2.15 : STATISTICAL ANALYSIS OF HOUSEHOLD WORK AND EDUCATION TRIP PRODUCTIONS

Trip purpose and Eqn.no.	Dependent variable	Mean dependent variable	Regression constant (a_o)	Standard error of estimate (SEE)	SEE as a percentage of mean dependent variable (SEEP)	Overall F-value	Critical F at 0.001 percent
Work trips							
1	TWM (all modes)	5202	7.34	73.87	1.42	19756.54	4.09
2	WM_1 (walk)	2481	70.76	221.02	8.91	321.59	4.00
3	WM_2 (bicycle)	2069	-23.76	200.50	9.69	441.46	4.39
4	WM_3 (motorcycle/ scooter)	318	-0.84	48.12	15.14	125.54	4.00
5	WM_4 (car)	39	-0.15	19.86	51.51	62.10	4.24
6	WM_5 (bus)	124	-31.01	73.80	59.38	18.55	3.75
7	WM_6 (cycle rikshaw)	171	-26.24	93.14	54.33	17.49	4.39
Education trips							
8	TEM (all modes)	3730	4.37	58.56	1.57	10370.74	3.32
9	EM_1 (walk)	2851	51.35	140.28	4.92	958.58	3.75
10	EM_2 (bicycle)	339	-7.69	107.29	31.69	40.00	4.39
11	EM_3 (motorcycle/ scooter)	13	0.79	11.11	-	9.47	4.00
12	EM_4 (car)	6	-0.05	8.22	-	9.89	4.24
13	EM_5 (bus)	159	-27.69	66.63	41.83	30.63	3.49
14	EM_6 (cycle rikshaw)	362	-5.94	96.29	26.59	39.22	4.00

3 MODE CHOICE DECISION MODELS

3.1 INTRODUCTION

A number of modal split models have been proposed using a variety of statistical techniques, including multiple regression, discriminant, probit and logit analysis (Warner, 1962; Lave, 1967; Lisco, 1967; Quaramby, 1967; Stopher, 1969). Logit model is closely related to economical and psychological theories of choice behaviour (Stopher and Meyburg, 1975). The models involve the comparison of evaluation measures which are linear combinations of the attributes of choice alternatives, and are compensatory in nature because they allow high level of satisfaction with one attribute to allow for low levels of satisfaction with others (Green and Wind, 1973). Besides the additive models, a number of nonlinear evaluation process models involving parabolic and hyperbolic response surfaces have also been formulated (Einhorn, 1970; Green and Wind, 1973).

To identify the behavioural mechanisms involved in human decision making process, a number of studies have been conducted by psychologists and market researchers. These are concerned with the development of choice models, which not only describe behaviour but also attempt to explain the processes by which choices are made (Kleinmuntz, 1968; Russ, 1972; Hughes, 1974a, b). Evaluation process research has produced a number of choice models which bear little resemblance to those derived from psychological and economic theory. These decision models provide more accurate prediction of consumer choices than simple addtive models (Forester, 1979).

The results obtained in the evaluation process research have, however, received very little attention from the transportation researchers. This study attempts to explore the use of models based on non-compensatory choice processes in modal choice. The non-compensatory choice processes are made on the basis of attribute by attribute comparisons of available alternatives and unlike compensatory structures they do not involve trade-off behaviour (Coombs, 1964; MacCrimmon, 1968). Only three choice models, namely, maximin, maximax and predominant choice models are explored in this study for mode choice decision process of the city of Kanpur. As the models are based on decision theory approach a brief review of decision theory is described in the following section.

3.2 DECISION THEORY

Decision theory is a study of how to make decisions under conditions of uncertainty (Jones, 1977). The decision maker is usually faced with number of uncertainties with respect to the nature of the environment. Basically the amount of uncertainty associated with a problem influences the approach used by the decision maker. If the state of nature is known with certainty for each action, the decision making process is called the decision under certainty. If only the probability of each state of nature for each action is known then the process is called decision under risk. If a course of action must be chosen without any knowledge of the state of nature, then this is decision under uncertainty.

Generally some type of information is known about the likelihood of the occurranceof each state of nature. But in the absence of any information about the states of nature, a number of decision procedures are available to use. In all the decision theory problems preperation of payoff table between the events and the actions is very important. From this payoff table a single numerical value is given to each event-action pair i.e., a row is assigned to each possible event that could conceivably occur. A column is assigned to each of the action alternatives under consideration. From the payoff table it becomes easy to take a decision. The possibility of adopting a single numerical value for each event-action pair is not a difficult task. A technique called utility theory allows to take many seemingly nonnumerical or multinumerical situations and develop for each event-action pair a single numerical quantity that represents the utility of that pair to the decision maker. Utility theory can also be used to convert multiple criteria into a single numerical criterion. Decision theory problems in general require:

(i) the decision maker has all potential actions known to him and he must choose exactly one of them;

(ii) a list of events (states of uncertainty) that are mutually exclusive and collectively exhaustive can be compiled;

(iii) to each event-action pair the decision maker must be able to assign a single numerical value to represent the consequence of choosing that action and having that event occur.

3.2.1 Concept of Utility

Utility theory is designed to take into account the magnitude of payoffs in decision problems as well as their expected value. It is assuming that the consequences of the actions being considered are serious enough so that the decision maker can not use expected value as a criterion for making his decision. The decision tree corresponding to the problem is then reduced, so that the decision maker is left with a choice between several chance outcomes. Therefore, ignoring the expected value criterion completely, the decision maker has still been able to develop a method for solving a decision problem with the assumption throughout that he is willing to accept risk. The type of utility described here is called cardinal utility. Economists also describe utility in a purely relational way which is called ordinal utility. This ordinal utility determines which of the two quantities is preferred to the other. This also considers the pairwise relationships between the two alternatives. The cardinal utility expresses not only that one is preferred to another but also by how much.

3.2.2 Methods of Selecting a Best Action

Maximin approach: The maximin approach operates under the assumption that the worst of all possible worlds is going to happen (Parsons,1974). It is designed to select the action alternative that will result in the most desirable consequence under the conditions of the action alternative that maximizes the minimum monetory payoff. The procedure starts first by

determining the least desirable consequence that can occur with respect to the action alternatives under consideration that is the minimum conceivable payoff. Then select as the optimal action the action alternative that would cause the least desirable consequence or minimum payoff to be maximized.

Maximax approach: This approach is based on extreme optimism. It operates under the assumption that best of all possible worlds will happen and is designed to select the action alternative that will generate the greatest payoff under the conditions of the action alternative that maximizes the maximum monetory payoff. The procedure starts first by determining the most desirable consequence that can occur for each action alternative under consideration, that is the maximum conceivable payoff. Then select as the optimal action, the action alternative that causes the most desirable consequence to be maximized i.e., maximizes the maximum payoff.

Selection of the action alternative with the maximum expected monetory payoff : This decision rule employed under the criterion of maximum expected monetory payoff is to select the action alternative that would generate the maximum average profit per period if that action alternative were to be pursued consistently over time. This criterion is most appropriate in repetitive situations, where the action selected will be implemented many times and where the magnitude of the most undesirable consequence is small.

3.3 MODE CHOICE MODELS FOR KANPUR

3.3.1 Introduction

Non-compensatory choices are made on the basis of attribute by attribute comparison of various alternatives. The levels of satisfaction associated with each of the attributes perceived by the person in the decision process are evaluated. These levels help in decision making. Based on the concept of decision theory it is decided to develop maximin, maximax and predominant-attribute models. The choice of attributes involved in the decision making process is of paramount importance. An individual in choosing the mode of travel evaluates some major attributes like travel time, travel cost etc., of the available modes of travel. The choice processes have to be inferred from observable phenomena. The following three approaches are generally used to make these inferences.

(i) Experimental approach: It involves the presentation of alternatives with the attribute levels determined by a full or fractional factorial design structure.

(ii) Analysis of trip makers survey data: In this approach the trip makers of different characteristics are asked to evaluate alternative modes of travel along with the reports of the actual mode choice.

(iii) Verbalised self reports from the decision makers: The reports are used to construct or evaluate simulation models of judgement process.

$$U_{i,j} = \underset{k}{\text{Min}} [U_{i,k}] \quad .. (3.1)$$

(ii) Selection of alternative: Identify the attribute 's' having maximum of the minimum utilities $[U_{i,j}]$ and select the alternative corresponding to that as the mode of travel

$$U_{s,j} = \underset{i}{\text{Max}} [U_{i,j}] \quad .. (3.2)$$

The maximin model involves the selection of alternative j with utility value $U_{s,j}$ such that

$$U_{s,j} > \underset{i}{\text{Max}} \ \underset{k}{\text{Min}} [U_{i,k}] \quad k \neq j \quad .. (3.3)$$

The decision process of the model is explained in Fig. 3.1. There are four attributes (1 to 4) and four alternative modes (1 to 4) . In the first step, the alternatives having the minimum utility levels for each attribute, i.e. $U_{1,4}$, $U_{2,2}$, $U_{3,1}$, and $U_{4,3}$ are identified. The maximum of the minimum utility level, i.e., $U_{4,3} = \text{Max} [U_{1,4}, U_{2,2}, U_{3,1}, U_{4,3}]$ is identified and the mode corresponding to $U_{4,3}$, i.e., mode 3 is selected.

3.3.3 Maximax Model

This strategy calls for the identification of the highest satisfactory alternative of each attribute and the selection of the alternative which yields the highest maximum level of satisfaction. Here also the process consists of two parts:

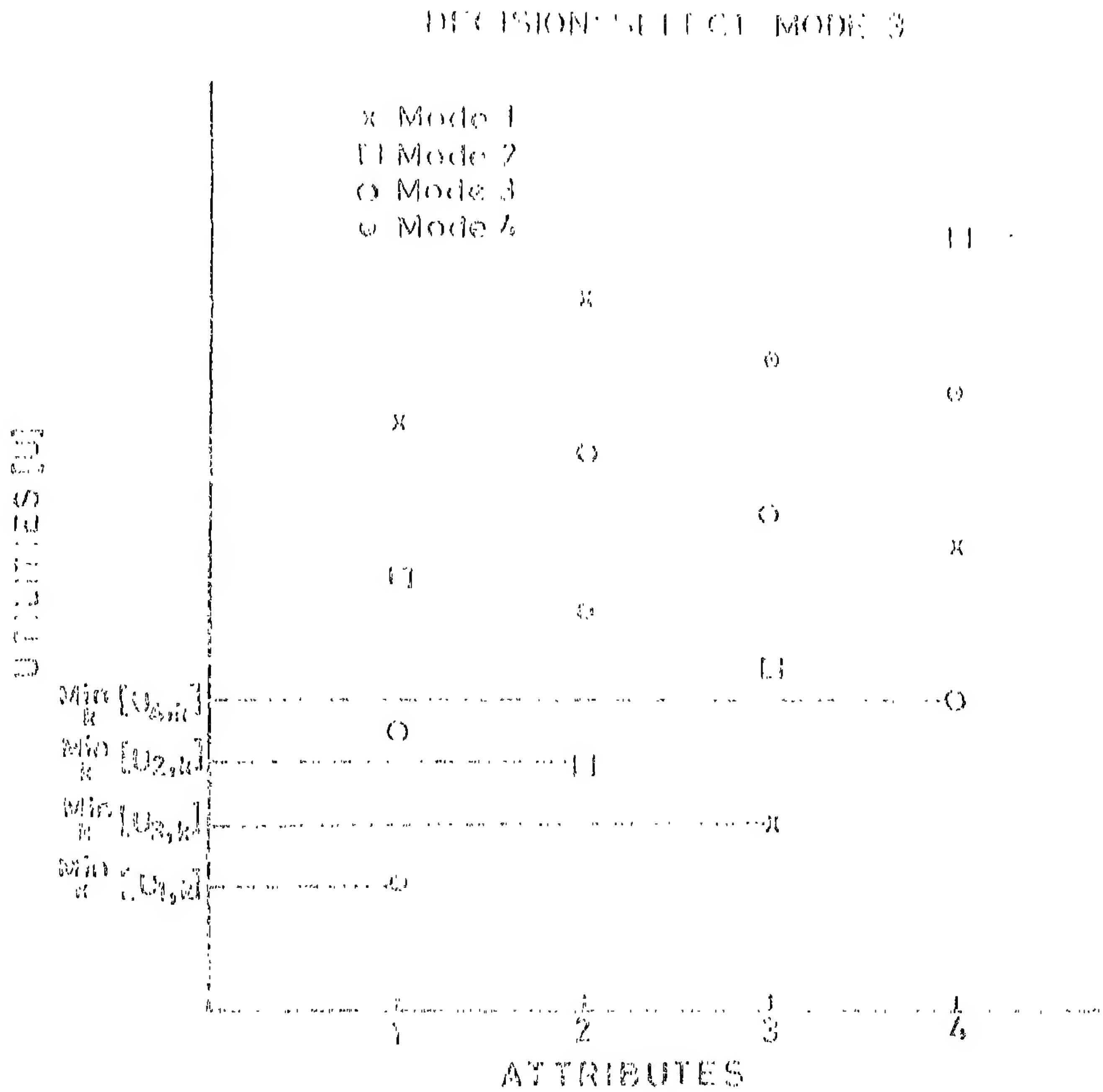


FIG.3-1 MAXIMIN MODEL

(i) Identification of alternative having maximum utility for each of the attribute: Using the decision profiles of various modes of travel and the attributes, the alternative having the maximum level of satisfaction in each attribute is identified. Let

$U_{i,l}$ = maximum utility of ith attribute that occur for lth alternative ,

$$U_{i,l} = \underset{k}{\mathrm{Max}} \, [U_{i,k}] \qquad .. \quad (3.4)$$

(ii) Selection of alternative: Identify the attribute 's' having maximum of the maximum utilities and select the alternative corresponding to that as the mode of travel

$$U_{s,l} = \underset{i}{\mathrm{Max}} \, [U_{i,l}] \qquad .. \quad (3.5)$$

The maximax model involves the selection of alternative l with utility value $U_{s,l}$ such that

$$U_{s,l} > \underset{i}{\mathrm{Max}} \; \underset{k}{\mathrm{Max}} \, [U_{i,k}] \quad k \neq l \qquad .. \quad (3.6)$$

The decision process of the model is explained in Fig. 3.2. Here also there are four attributes with four alternative modes . In the first step the alternatives having the maximum utility levels for each attribute i.e., $U_{1,4}$, $U_{2,2}$, $U_{3,1}$ and $U_{4,2}$ are identified. The maximum of the maximum utility level, i.e., $U_{2,2} = \mathrm{Max} \, [U_{1,4}, U_{2,2}, U_{3,1}, U_{4,2}]$ is identified and the mode corresponding to $U_{2,2}$, i.e.,mode 2 is selectined.

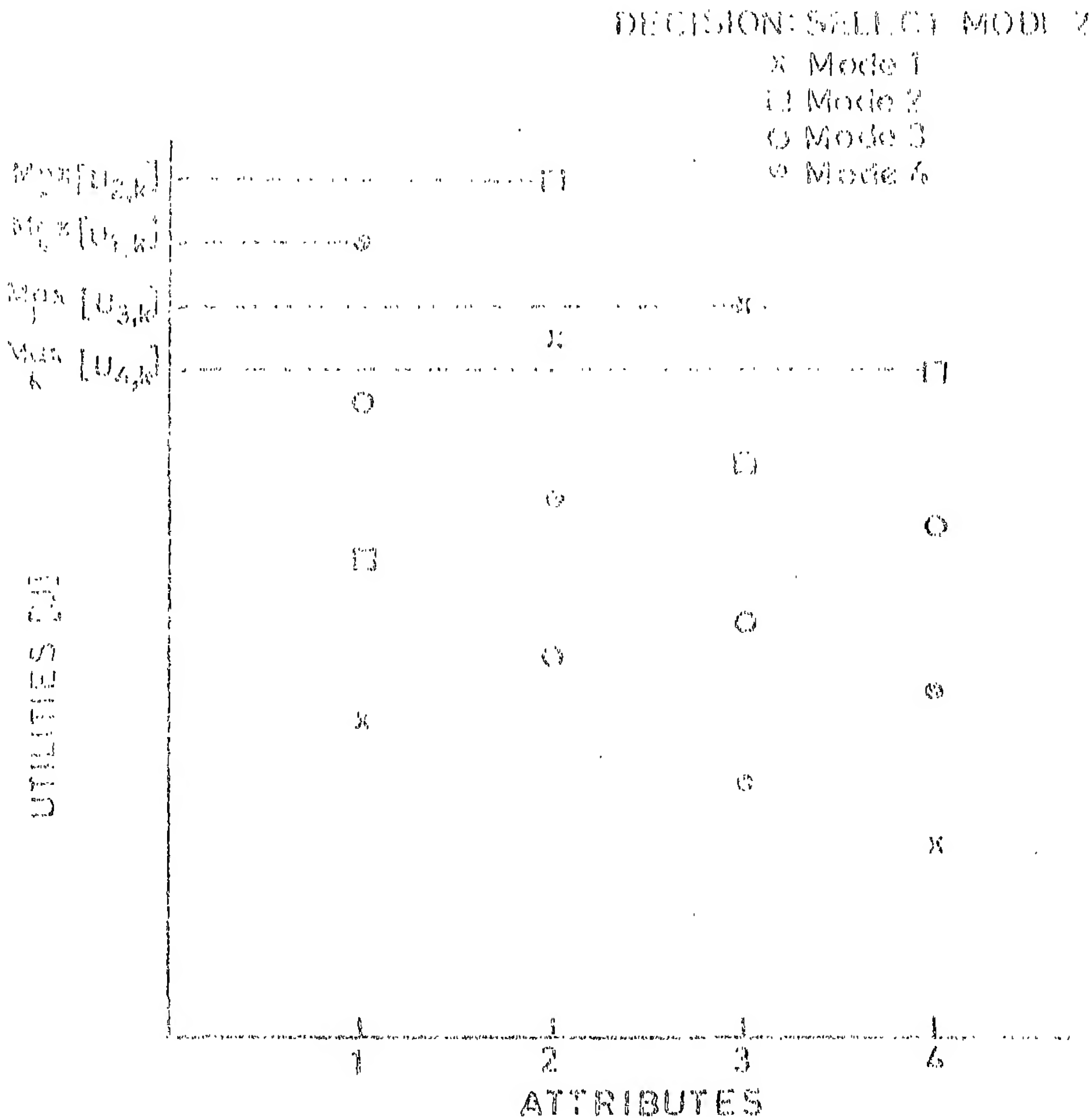


FIG.3.2 MAXIMAX MODEL

3.3.4 Predominant-Attribute Model

In this model the attributes are hierarchically ordered according to their generic importance. This is done by using the relative utilities of each of the attributes for an alternative. The relative utility value for an attribute i of an alternative k, $U'_{i,k}$ is expressed as

$$U'_{i,k} = \frac{U_{i,k}}{\sum_{i} U_{i,k}} \qquad .. \quad (3.7)$$

For each alternative the most predominant attribute is being identified based on utility level. In some cases the predominant attribute may not be same for each of the alternative , i.e., different alternatives may have different predominant attributes. For alternative modes of walk and bicycle, income may be the most predominant attribute, whereas for travel by car or motor-cycle/scooter the most predominant attribute is the ownership of the vehicle. The decision process consists of two parts:

(i) Select the most predominant attribute 'p' for each of the alternatives k , ($k = 1,2,\ldots n_m$) as the one with the maximum level of satisfaction, i.e.,

$$U'_{p,k} = \operatorname*{Max}_{i} \left[U'_{i,k} \right] \qquad .. \quad (3.8)$$

where $U'_{p,k}$ = relative utility of the most predominant attribute of kth alternative. This is obtained from the decision profile of various attributes for each alternative.

(ii) Selection of alternative: Select the alternative j having the maximum of the utility corresponding to the

most predominant attribute $U'_{p,k}$ for each alternative, i.e.

$$U'_{p,j} = \underset{k}{\mathrm{Max}} \left[U'_{p,k} \right] \quad .. \quad (3.9)$$

The decision process of the model is explained in Fig. 3.3. In the first step the attributes having the maximum utility level for each of the alternatives i.e., $U'_{3,1}$, $U'_{4,2}$, $U'_{1,3}$, and $U'_{3,4}$ are identified. The maximum of the most predominant attribute, i.e., $\mathrm{Max}[U'_{3,1}, U'_{4,2}, U'_{1,3}, U'_{3,4}]$ is identified and the corresponding mode $U'_{3,4}$, i.e., mode 4 is selected.

3.3.5 Model Inputs

(A) Choice of attributes

For the non-compensatory decision models formulated in sub-suctions 3.3.2. to 3.3.4, the choice of attributes for different modes is highly significant before estimating the level of satisfaction associated with each of the attributes. Based on the available information of mode choice process for Kanpur city, the following four attributes are selected for the decision process:

(i) family size
(ii) income level
(iii) vehicle ownership
(iv) trip length

The attributes (i) and (ii) represent the socio-economic standard of the individual. An individual with high income level may opt for a private motor transport (motorcycle/scooter, car etc.) or an expensive public transport. For longer trips the available alternatives are few compared to those available for shorter trip

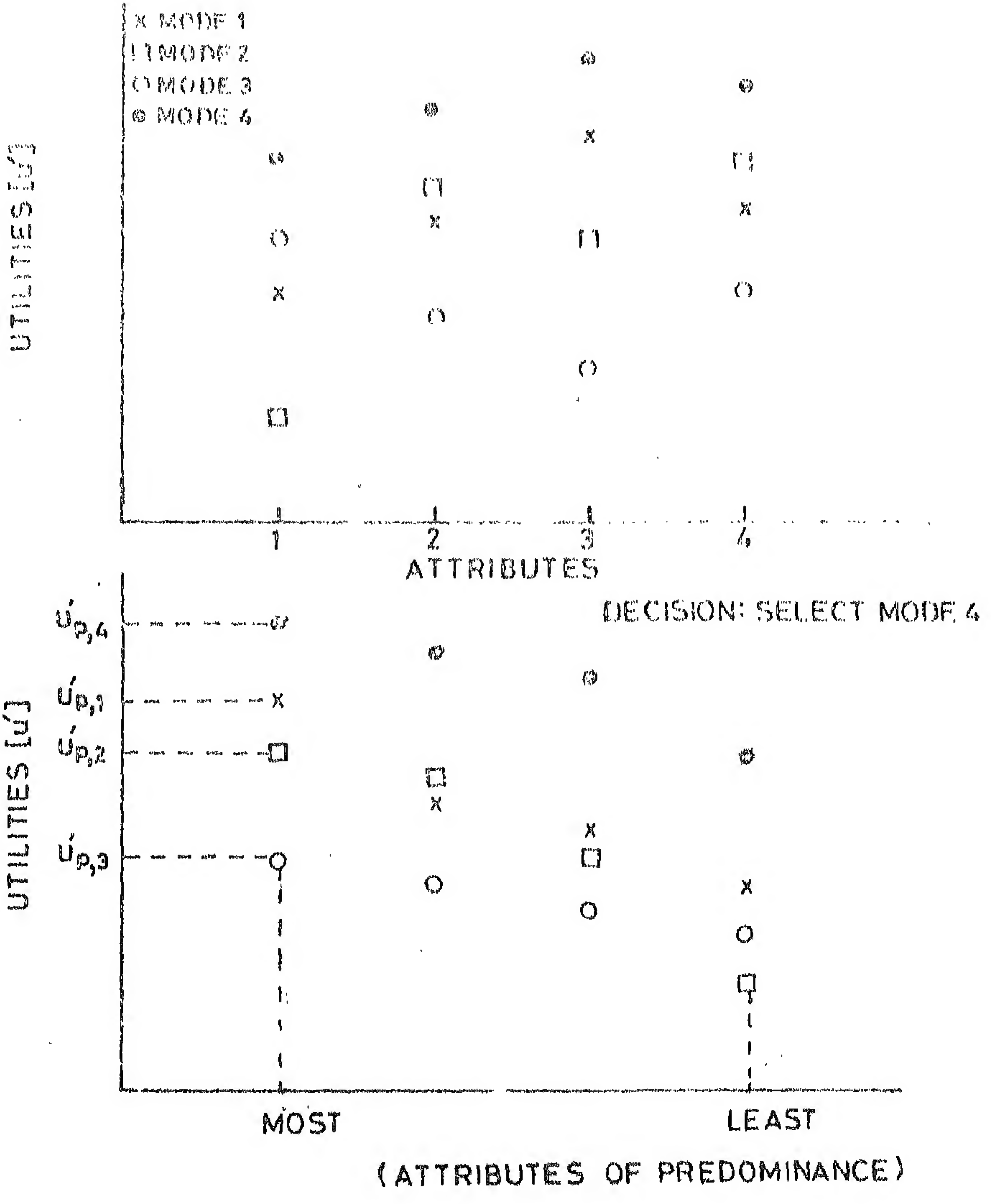


FIG. 3·3 PREDOMINANT ATTRIBUTE MODEL

lengths. The decision making process is further highly dependent on the type of vehicle ownership and the availability of public transport system . The level of satisfaction associated with each of the above attributes are estimated from the household survey data as explained in Section 2.3. Each of the above attributes are divided into number of different subgroups to classify an individual. These subgroups are given below :

(i) Family size

- family upto 4 members
- family with 5 members
- family with 6 members
- family with more than 6 members

(ii) Income level

- low income group
- lower middle income group
- upper middle income group
- high income group

(iii) Vehicle ownership: The mode choice decision process depends on the types of vehicles owned by the family. Eight subgroups of vehicle ownership as identified in this study are:

- no vehicle
- bicycle
- motorcycle/scooter
- car
- bicycle and motorcycle/scooter
- bicycle and car
- motorcycle/scooter and car
- bicycle, motorcycle/scooter and car

In situations when public transit system is available for travel the mode choice decision process may be different than when it is not available. Two broad groups that have been identified are

(a) accessibility to public transit system

(b) no accessibility to public transit system

In each of the above groups all the eight subgroups of vehicle ownership are included. There appears to be a significant need to incorporate the accessibility to public transit system as all zones of a typical Indian city may not be covered by the public transit system. The complete sub-groups of vehicle ownership attributes as identified in the decision process models are

(a) accessibility to public transit system and one of the following vehicle ownership group:
- no vehicle
- bicycle
- motorcycle/scooter
- car
- bicycle and motorcycle/scooter
- bicycle and car
- motorcycle/scooter and car
- bicycle, motorcycle/scooter and car

(b) no accessibility to public transit system and one of the following vehicle ownership group:
- no vehicle
- bicycle
- motorcycle/scooter

- car
- bicycle and motorcycle/scooter
- bicycle and car
- motorcycle/scooter and car
- bicycle, motorcycle/scooter and car

(iv) Trip length

- upto 1 kilometer
- 1-2 kilometers
- 2-3 kilometers
- 3-5 kilometers
- more than 5 kilometers.

In addition to the selected attributes, the location of residence of an individual also can be included.

(B) Estimation of utility measure

The formulated models call for estimation of the utility for each attribute of the various alternatives. Non-availability of appropriate data for Kanpur restricted the estimation of utility measure in terms of time or cost values. In this study, the utility level of an attribute for a particular mode is expressed as a proportion of the trips made by that alternative out of the total trips of the concerned attribute. The relative utilities of the different attributes in an alternative $U'_{i,k}$ are also estimated from the total trips in the various cells of concerned alternative. The household survey data is aggregated for the whole city and the utility measures of each of the 29 subgroups of attributes are estimated for all the modes of travel. Survey indicated that there are in general 6 modes of travel, namely, walk, bicycle, motorcycle/scooter, car,

bus and cycle rikshaw. A major share of these trips are made on foot within the CBD area over short distances. The non-compensatory mode choice decision models are tested only for 5 modes of travel in which some vehicle is used and the trips on foot are not considered for estimation of utility measures and analysis.

3.3.6 Model Outputs

The full factorial design of the decision process involves 4x4x16x5 = 1280 decision cells. Each value indicates the number of subgroups for an attribute. The choice of mode for each of the cells is made seperately by the three decision making processes of maximin, maximax and predominant-attribute models. The number of trip makers in each of the decision cells are identified and the total trips by each of the alternative modes are determined. Analysis is carried out for 5 percent sample size. The number of trips to be performed by the alternative modes as per the three models along with the observed trips are given in Table 3.1 for work trips and Table 3.2 for the education trips. The results obtained from maximin and maximax models compare well with the observed data, but the results obtained from the predominant-attribute model significantly vary from the observed values. As such the most predominant-attribute model as formulated does not appear to be suitable for the city of Kanpur. This may be attributed to the fact that the utility value for each attribute is taken as the proportion of trips for that alternative, or it may be that the mode choice decision process is not affected only by the

TABLE 3.1 : OBSERVED AND MODEL TRIPS BY DIFFERENT MODES OF TRAVEL FOR WORK TRIPS

Type of mode	Observed trips (O)	Trips by maximin model (M_1)	Percentage difference ($\frac{M_1-O}{O}$x100)	Trips by maximax model (M_2)	Percentage difference ($\frac{M_2-O}{O}$x100)	Trips by predominant-attribute model (P)
Bicycle	4967	5092	+ 2.52	4956	- 0.22	625
Motorcycle/ scooter	762	650	-14.69	782	+ 2.62	1226
Car	93	80	-13.98	84	- 9.67	990
Bus	298	298	0.00	298	0.00	2751
Cycle rikshaw	412	412	0.00	412	0.00	940
Mean percentage difference		=	-5.23		-1.45	
Mean absolute percentage difference		=	6.24		2.50	

most predominant attribute. It is desirable to test this model for some other cities also to identify the appropriate cause for differences.

The results by maximin model indicate that the highest variation between model and observed values is 14.69 percent for the trips performed by motorcycle/scooter whereas for bicycle trips the variation is only 2.52 percent. Another highly important observation is that the model trip values obtained for bus and cycle rikshaw exactly match with the observed values. The mean absolute percentage difference is 6.24. The results obtained from maximax model improve upon those by maximin model the maximum variation being 9.67 percent for the trips performed by cars whereas the variation is less than 3.0 percent for other cases.

The comparison of the trips obtained from the models with the observed values as given in Table 3.2 for education trips indicate that the maximin and maximax models give quite sensible results, in this case also, whereas the predominant-attribute model trips significantly vary from the observed values. The maximum variation in maximin model output is 7.69 percent for the trips performed by car whereas for bicycles it is only 4.43 percent. An important observation is that the value of bus trips obtained from the two models exactly matches with the observed ones, whereas in case of cycle rikshaw trips, the variation is 4.49 percent for maximin model. The results obtained from the maximax model show that the maximum variation is 16.12 percent for the trips performed by

TABLE 3.2 : OBSERVED AND MODEL TRIPS BY DIFFERENT MODES OF TRAVEL FOR EDUCATION TRIPS

Type of mode	Observed trips (0)	Trips by maximum model (M_1)	Percentage difference ($\frac{M_1-0}{0}$x 100)	Trips by maximax model (M_2)	Percentage difference ($\frac{M_1-0}{0}$x100)	Trips by predominant-attribute model(P)
Bicycle	812	776	- 4.43	805	- 0.86	18
Motorcycle/ scooter	31	29	- 6.45	26	-16.12	93
Car	13	12	- 7.69	11	-15.38	1005
Bus	383	383	0.00	383	0.00	840
Cycle rikshaw	869	908	+ 4.49	883	+ 1.61	152
Mean percentage difference =			- 2.82		- 6.15	
Mean absolute percentage difference =			4.61		6.79	

motorcycle/scooter whereas the variation is only 1.0 percent for bicycle trips. The mean absolute percentage difference is 6.79.

The results explained above indicate that the choice models are able to explain the decision process. A major share of the trips are made by slow moving vehicles like bicycle andcycle rikshaw. To test the choice amongst the motor vehicular modes, experiments are also conducted on the models by considering only three alternative modes namely motorcycle/scooter, car, and bus. The number of trips to be performed by these models are compared with the observed values and are given in Table 3.3 for work trips and in Table 3.4 for education trips. It may be mentioned that only maximin and maximax models are used for decision making process in these cases. The results indicate that the model values are very close to the observed ones in these cases also. The variation in fact is less than observed from the previous case when 5 alternatives are considered.

The results of the study indicate that non-compensatory decision choice models have a bright future to represent the mode choice decision process. The aim of this study is of only exploratory nature. It is desirable to test a broad range of choice models before adopting one for mode choice behaviour. Additional research in the choice of modal attributes involved in travel decisions is also called for.

TABLE 3.3 : OBSERVED AND MODEL TRIPS WHEN MOTOR VEHICULAR MODES ARE CONSIDERED FOR WORK TRIPS

Type of mode	Observed trips (O)	Trips by maximin model (M_1)	Percentage difference ($\frac{M_1-O}{O}$x100)	Trips by aximax model (M_2)	Percentage difference ($\frac{M_1-O}{O}$ x100)
Motor-cycle/scooter	762	770	+ 1.04	766	+ 0.52
Car	93	85	- 8.60	89	- 4.30
Bus	298	298	0.00	298	0.00
Mean percentage difference			= - 2.52		- 1.27
Mean absolute percentage difference			= 3.20		1.61

TABLE 3.4 : OBSERVED AND MODEL TRIPS WHEN MOTOR VEHICULAR MODES ARE CONSIDERED FOR EDUCATION TRIPS

Type of mode	Observed trips (O)	Trips by maximin model (M_1)	Percentage difference ($\frac{M_1-O}{O}$x100)	Trips by maximax model (M_2)	Percentage difference ($\frac{M_2-O}{O}$x100)
Motor-cycle/scooter	31	32	+ 3.23	32	+ 3.23
Car	13	12	- 7.69	12	- 7.69
Bus	383	383	0.00	383	0.00
Mean percentage difference			= - 1.49		- 1.49
Mean absolute percentage difference			= 3.64		3.64

4 SIMULATION MODEL FOR COMBINED MODAL SPLIT AND TRIP DISTRIBUTION

4.1 DISAGGREGATE MODELS

Most of the conventional forecasting models were generally conceived using analogies with physical systems with traffic flows described in terms of hydraulic or gravity flow models. Different model components were not developed from a unified framework. For example, a trip generation model may be developed independent of a modal split model. These conventional models often produce predictions which are counter intutive . For example , trip generation models take no account of network characteristics while the trip distribution models have only travel time as the service attribute (Binder , 1973; Bouchard, 1973; Cambridge Systematics Inc.,1974). These models are generally estimated by fitting relationships to the aggregated data and express aggregate correlations that are not based on a logical behavioural specification. Inspite of involving costly and time consuming data gathering and computational requirements, the conventional models are not easily adopted to short run planning and transportation system management (McFadden, 1978). Most of the aggregate models cannot be used to explain why a change has happened or to predict what will happen in the future. Observed aggregate correlations do not adequately explain the underlying cause and effect relationships.

Travel demand is a complex behavioural process. To produce valid policy sensitive forecasts brhavioural model based

on sound theories and valid statistical procedures are required. The basic concern of the travel demand modelling is the trip. Each trip has an origin and destination. It also involves the number and characteristics of each of the traveller's mode, route, frequency, destination, time of travel, purpose etc. (Warner, 1962; Lisco, 1967; Lave, 1967 ; McGillivray,1972; Charles River Associates, 1972; Peat, MarWick and Mitchell,1973). The process of aggregation inherently losesthis information. Travel is a very disaggregate process and therefore raw data concerning travel is typically disaggregate. Substantial progress has been made to develop and implement travel demand forecasting methods which are behaviourally sound and policy relevant. The resulting behavioural disaggregate methods expand the policy sensitivity of forecasts. These methods are based on a unified conceptual framework and travel demand is generated by individual choice behaviour and more specifically generated by maximization of preferences or utilities (McFadden, 1978).

Disaggregate forecasting is an approach or system for building models and it can provide the planner with the method of dealing with a variety of problems as they occur. A major advantage of disaggregate models is that they allow the planner to address questions such as the demand for a new mode, which are otherwise difficult to answer in a conventional framework. In disaggregate behavioural choice theory the individual is the basic decision making unit choosing from available alternatives the most desirable. The choice depends upon the characteristics

of the individual. A complete definition of a transportation alternative for an individual includes total pattern of travel, location of residence and job, number and type of vehicles owned, frequency of different trip purposes, scheduling of trips, mode choice, and route choice. Travel demand models generally concentrate on certain dimensions of travel behaviour. Disaggregate models are typically probabilistic over the choices addressed by the models (Ben-Akiva, Lerman, Manheim, 1976), i.e., each choice is predicted with some probability instead of with certainty. Since the models consider invidual decision maker, the distribution of the impact of a given policy can be determined. In summary, a disaggregate behavioural model is specified by forming a concrete individual utility function, a probability distribution of the unobserved variables and a share of each market segment in the population.

The use of disaggregate models is more flexible and adoptable than conventional aggregate models; they are easier to use and can be used in more applications and in ways that are more policy responsive (Lerman, 1975; Lerman and Ben-Akiva, 1975). Models that have been initially developed for urban passenger planning related to work place choice, mode to work, mode for non-work trips, frequency for non-work trips; destination choice for non-work trips, and time of day for non-work trips (Paul O. Roberts, Jr. 1978). Some models which can be used to make joint or simultaneous choices related to joint auto ownership/mode choice to work, joint frequency/destination/mode choice for shopping and sociorecreational travel (Ben-Akiva, 1974), joint frequency/destination/mode choice for week end

travel (Lerman, 1975), and joint residential location/housing type/auto ownership/choice of mode to work(Ben-Akiva and Adler, 1975).

A system of models have been developed in a joint research of Massachusetts Institute of Technology and Cambridge Systematics, Inc. (Ben-Akiva, Lerman and Manheim, 1976). The models developed in this research share a common set of characteristics both in their theoretical understanding and their implementation. The significant commonalities are individual choices, explicit behavioural theory, explicit treatment of multiple dimensions of choice, and valid estimation methods. The disaggregate models have also been applied for a number of cities of U.S.A. The disaggregate modelling approach is issue oriented, easily tailored to the problem, flexible in use of both short-range as well as long-range planning problems, and policy sensitive.

4.2 CHOICE OF SIMULATION FOR DECISION MAKING PROCESS

A travel demand model is concerned with decisions which result in trip making. These decisions are at household and individual levels. There are some other choices which are interrelated with actual trip making choices and it is difficult to seperate the two. For example the choice of residential location is not a trip making decision in itself. But the combination of a workers employment location choice and his/her residential location decision has as its consequence a daily work trip. The general framework from which the disaggregate models can be derived should begin with the partitioning

of all household and individual decisions into two sets, viz., those relevant to transportation analysis and those which for practical purposes can be ignored. This partitioning will produce the following vector of household decisions (Ben-Akiva, Lerman and Manheim, 1976), employment location, residential location, housing type, automobile ownership, mode to work, frequency, destination, time of day, mode to non-work and route.

The above factors will describe the decisions, which the complete model system must consider. This perspective, however, would produce a model of unmanageable dimensions, since the entire vector of possible choices would be limitless. It would thus be impossible to develop useful models. There are, however, some interrelationships among components of the above vector which are of different character than others. Some of the decisions like residential location choices are stable over fairly long periods while other choices such as social/shopping trip frequency are altered almost on daily basis. Some decisions can be logically represented as being made collectively by the household while others can be of individuals choice. It may thus be possible to formulate explicit behavioural hypotheses. This helps to establish a structure of the total vector of choices as a logical working hypothesis.

Different designs for disaggregate models have been suggested by various researchers to solve complex interaction processes, but the implementation has been highly hampered by

the methodological difficulties. The disaggregate methods cannot treat large amount of data nor do they offer a reasonable logic to produce decision processes on the individual level in a complex travel system (Kreibich, 197). In India , the travel behaviour is highly complex because of wide variations in the modal characteristics and also large variations in the socio-economic status of the individuals. Simulation of this complex decision process at individual level is attempted in this study. In the proposed model the modal choice and the trip end decisions of the individuals are met in an integrated framework. The mode choice and trip distributions obtained from the simulation model are tested with the observed travel data for Kanpur metropolis. The details of the model are given in the following sections.

4.3 CONCEPT OF STATUS GROUPS

The concept of status groups is used in the proposed simulation model because of the importance of mobility constraints. Individuals with similar constellations of activity constraints are grouped into status groups. All the members of a status group are exposed to similar internal and external decision factors. In addition, it provides a framework to consider demographic and socio-economic as well as locational determinants of transportation behaviour simultaneously. In a socially oriented planning framework they can be considered as the target groups of planning concepts and measures. The proposed concept of status groups can be

considered as a first operational definition of constraint oriented approach to household travel behaviour. The indicator variables that can be used for defining the status groups in a typical major Indian city are:

(i) for work trips

- residential location
- household income

(ii) for education trips

- residential location
- household income
- level of education

The choice of the three indicator variables for education trips results in a large number of decision situations. To simplify it the household income is not considered as an indicator variable of education status group for the city of Kanpur. Besides the residential location, the level of education (primary, secondary, degree, vocational) is thought to be more significant than the household income.

For each status group four dependent variables which are considered to be of strategic importance for the decision process of mode choice and trip end choice are:

(i) model of travel

(ii) employment/education opportunities in a zone

(iii) travel time

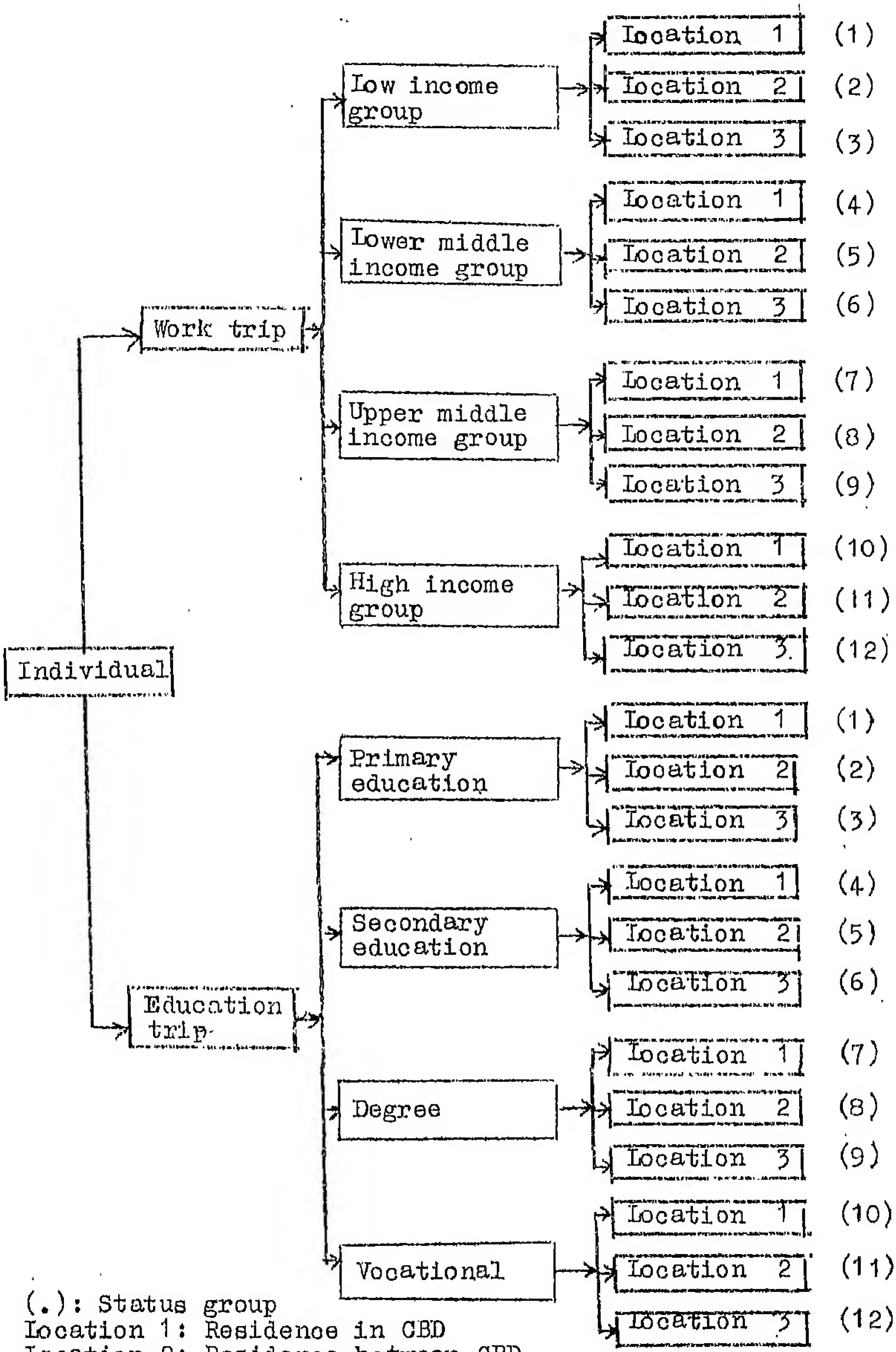
(iv) travel cost.

Besides these four dependent variables for each status group, like travel time and travel cost for each of the mode of travel independent of status groups are also considered to be significant importance for the decision process. The decision profiles for each of the dependent variable of a status group are derived from the available travel survey data. These profiles determine the behaviour of an individual within a stochastic framework.

4.4 DECISION PROCESS INPUTS

4.4.1 Status Group of Individual

Simple cross tabulation measures of individual travel behaviour by demographic and socio-economic variables reveal distinct differences with respect to network patterns and travel times. A constraint oriented approach using the concept of status groups requires a multivariate design. The method of cluster analysis is therefore applied to group the individuals into status groups with homogeneous decision situations. The pattern of constraints which determine travel behaviour of status groups points out the importance of stage in the family life cycle and of resulting household linkages. Fig. 4.1 shows how the individuals of a zone are classified into different status groups for work and education trips. As mentioned in Section 4.3 for work trips the status groups are decided based on residential location and household income. The household income is classified into four groups such as low income, lower middle income, upper middle income, and high income. For each income group the status groups are



(.): Status group
Location 1: Residence in CBD
Location 2: Residence between CBD and intermediate cordon
Location 3: Residence between intermediate and outer cordons

FIG. 4.1 : STATUS GROUP OF AN INDIVIDUAL

decided based on residential location such as the residence in the CBD area, between CBD area and intermediate cordon, and between intermediate and outer cordons. For education trips the status groups are decided based on level of education and residential location. The level of education is classified into four groups, namely, primary, secondary, degree, and vocational. For each group based on residential location the status groups are decided. There are 12 status groups for work trips and the same number for education trips for the city of Kanpur.

4.4.2 Distributions of Employment/Education Opportunities

The model as formulated in this study simulates the decision process of the individual. The decision profiles of the dependent variables considered for the decision process are to be developed for each status group. The city under consideration is partitioned into three cordons namely, central (CBD), intermediate, and outer. The area between the cordons is divided into zones such that each zone is homogeneous with regard to population density, and land use pattern. The boundary of a zone is also marked based on the location of the streets in the network. The household survey data is classified by purpose into various zones and status groups. The number of attractions by purpose to a zone of each status group is determined from the household survey data. Based on the number of attractions of each status group in a zone, all the zones of the survey area are classified into certain groups. The zones within each of these groups specifies equal employment/education facilities available for an individual of a certain status group. The procedure for this classification is shown in Fig. 4.2.

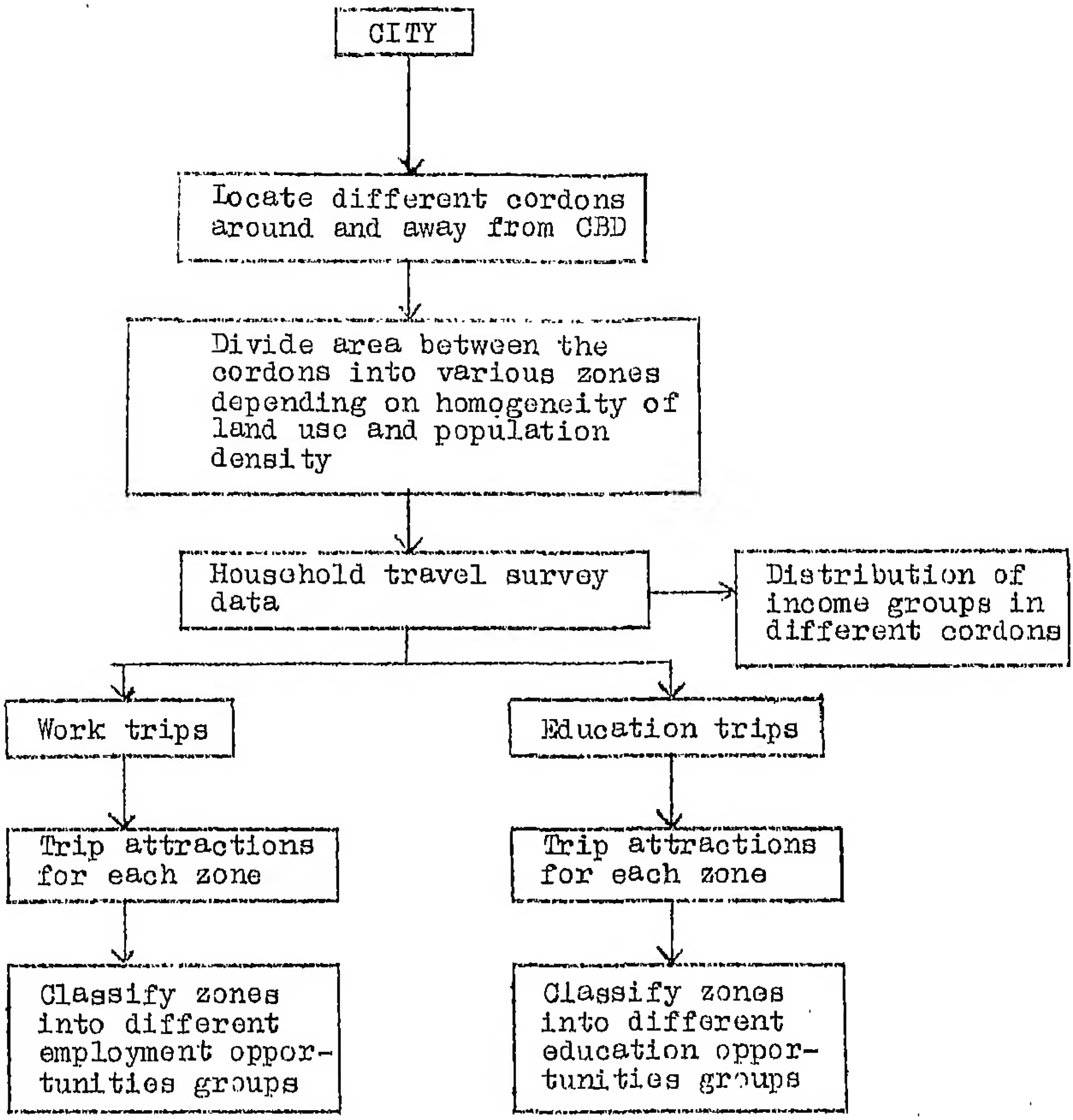


FIG. 4.2 : DISTRIBUTIONS OF EMPLOYMENT/EDUCATION OPPORTUNITIES

4.4.3 Decision Profiles

Four variables considered to be of strategic importance for the decision process of an individual are mode of travel, employment/education opportunities, travel time, and travel cost. The decision profiles of these dependent variables for each of the status groups need to be determined from the household survey data, and the procedure is shown in Fig.4.3. The procedure for determining the distribution of the different status groups is also shown in Fig. 4.3. The decision profiles represent major constraints on travel behaviour for each status group. These decision profiles are considered as the dependent variables and computed as frequency distributions. When the model is used to determine the effects of alternative concepts of travel behaviour, the decision profiles could also be derived from assumptions based on theoretical concepts. Accordingly the distribution of status groups in the origin zones would depend on the population distribution of a certain settlement policy, if it is to be tested with respect to its effects on the transportation system. The distribution of individuals by status group among origin zones in the study area reflects the influence of the transportation system on their decision profiles. The decision profiles which are being derived from the data are as follows:

- mode of travel
- employment/education opportunities
- travel time
- travel cost.

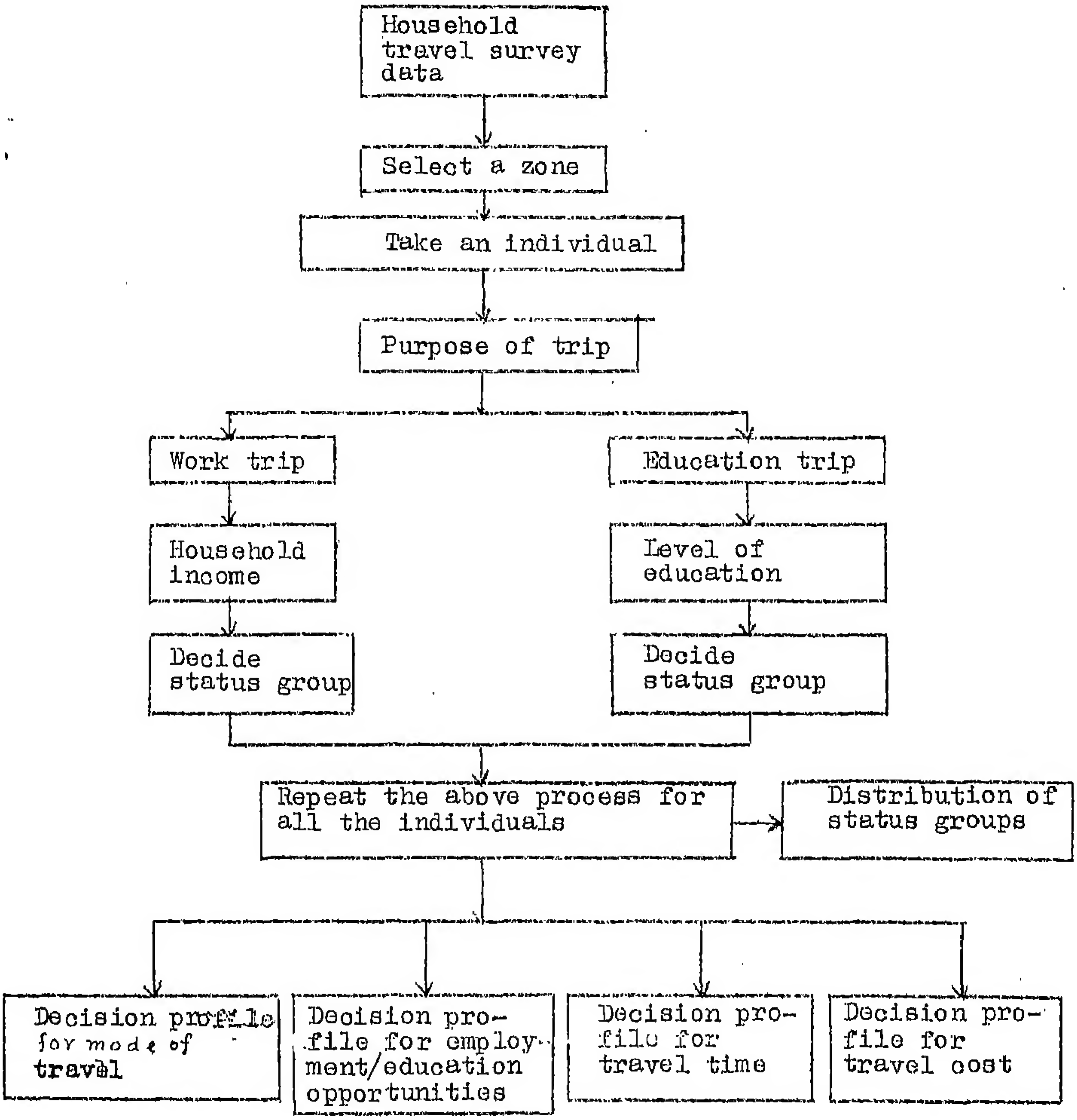


FIG. 4.3: PROCEDURE FOR DETERMINING DECISION PROFILES

(i) Mode of travel

The decision profiles for the mode of travel of each status group are derived from the household survey data. Six modes of travel, both public and private, including walk are considered. Twelve decision profiles of each status group for work trips and an equal number for education trips are derived. A representative decision profile of status group no.8 both for work and education trips are given in Fig. 4.4. This profile gives the frequency in percentage of trips performed by each of the modes of a particular status group. While generating the mode of travel for an individual of a particular status group, the cumulative frequency distributions are used.

(ii) Employment/Education opportunities

From the household survey data the number of trips by purpose attracted to each zone are determined. Based on these attractions the zones are classified into ten groups, each group representing a certain range of work/education opportunities available for a zone. The opportunities available for the ten groups are upto 1000, 1001-2000, 2001-4000, 4001-6000, 6001-8000, 8001-10000, 10001-12000, 12001-14000, 14001-16000, and more than 16000. The frequency distributions of the trips by purpose attracted to each of the above ten groups of zones are shown in Fig. 4.5. This decision profile is for one particular status group and similar ones are derived for each of the status groups by purpose. Using the cumulative frequency

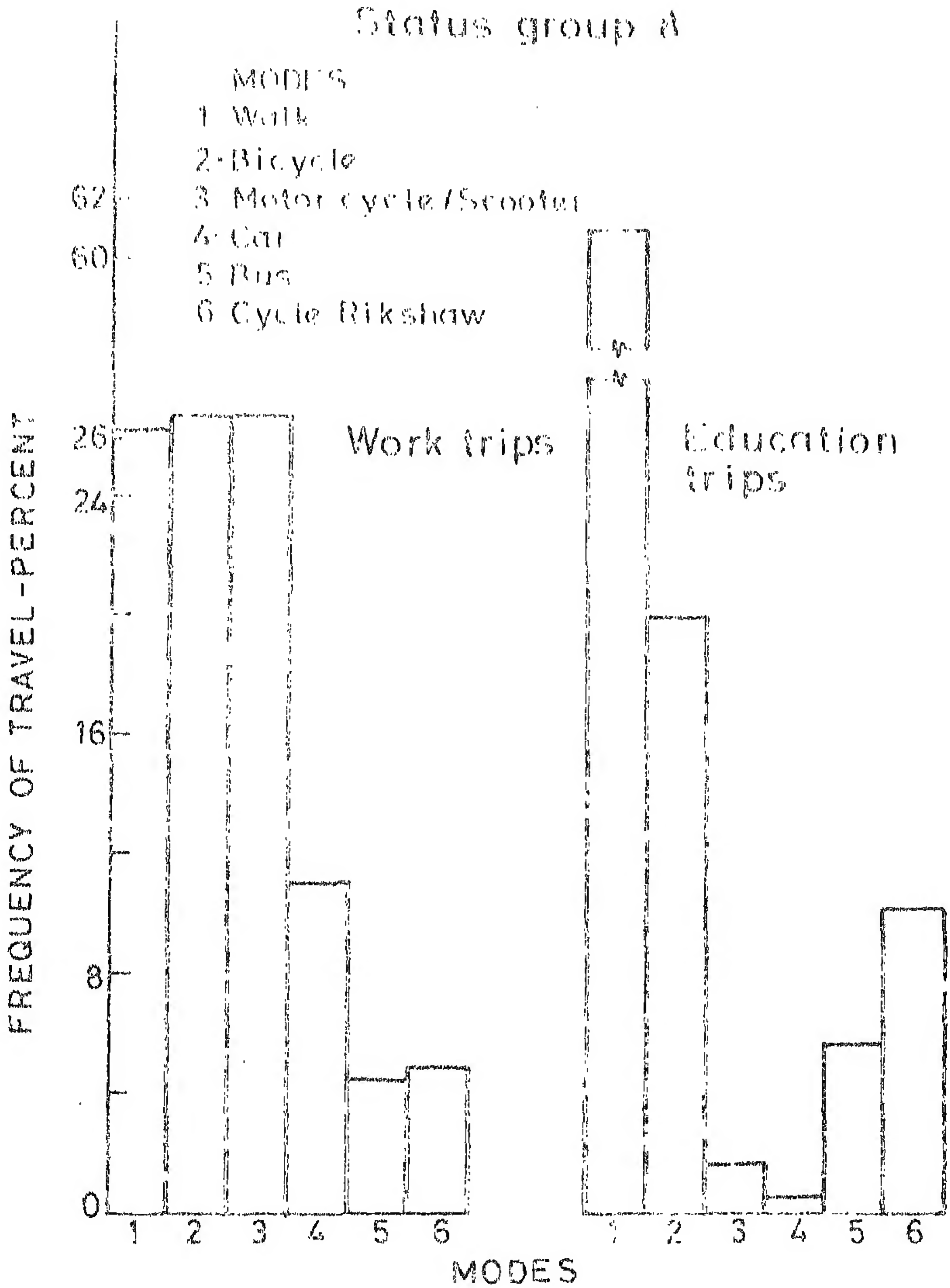


FIG.4.4 DECISION PROFILE FOR TRAVEL MODES

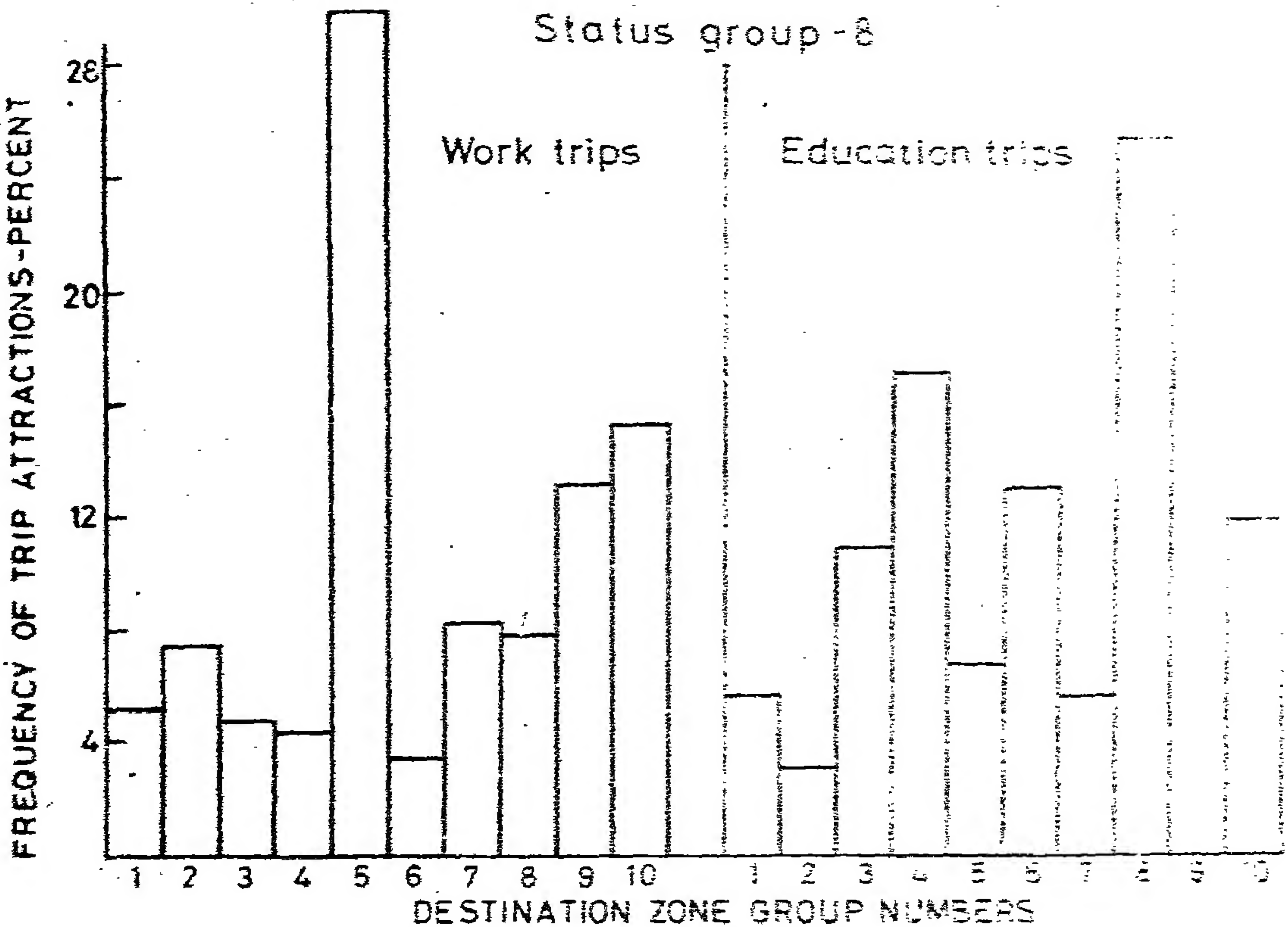


FIG-4-5 DECISION PROFILE FOR EMPLOYMENT/EDUCATION OPPORTUNITIES

distribution of these decision profiles, the model generates one of the ten groups of zones to which the trip may be attracted. This further helps in selection of the actual zone of destination.

(iii) Travel time

Based on the travel time of the observed trips from household data, the travel times are stratified into ten groups, each represented by a 5 minute interval. For a particular status group the frequency distributions of each of the travel time groups are represented in Fig. 4.6 for work and education trips. Similar decision profiles are derived for each of the status groups by purpose. It may be mentioned that these pro - files are aggregated for all modes of travel and stratification is only for status group type. The travel times of the modes available in Kanpur vary significantly, some being very slow like walk or by human driven vehicles whereas fast modes also are available. To properly stratify the modes of travel, decision profiles of travel time for each of the modes are also derived, and are shown in Fig. 4.7 for a particular mode. It may be worth mentioning that these decision profiles are independent of the status group, i.e., individuals of all status groups are aggregated.

(iv) Travel cost

Knowing the length and mode of the observed trips, their travel costs are estimated based on the unit operating cost of different vehicles. The travel costs by bus are as

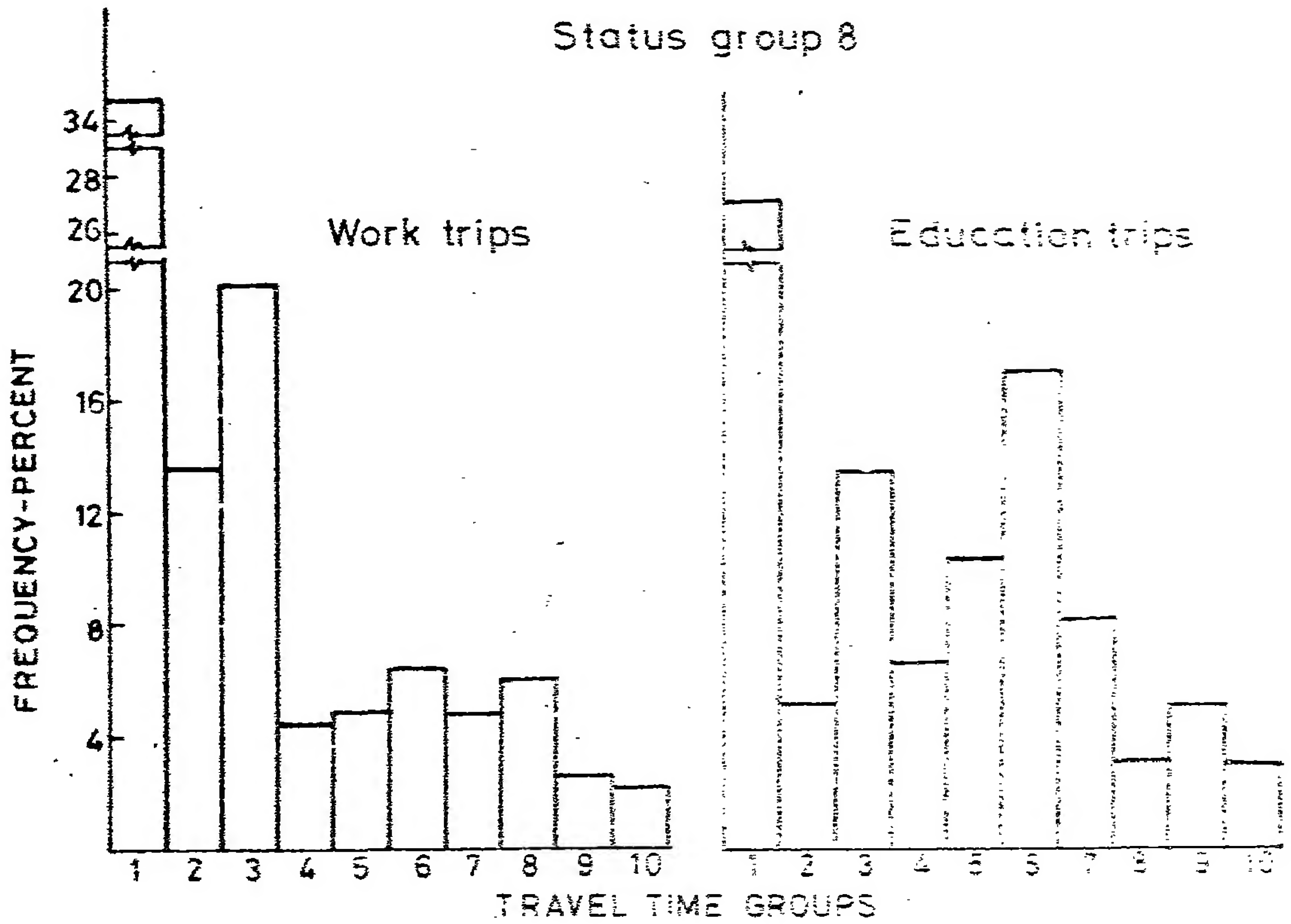


FIG. 4.6 DECISION PROFILE OF TRAVEL TIMES FOR STATUS GROUP 8

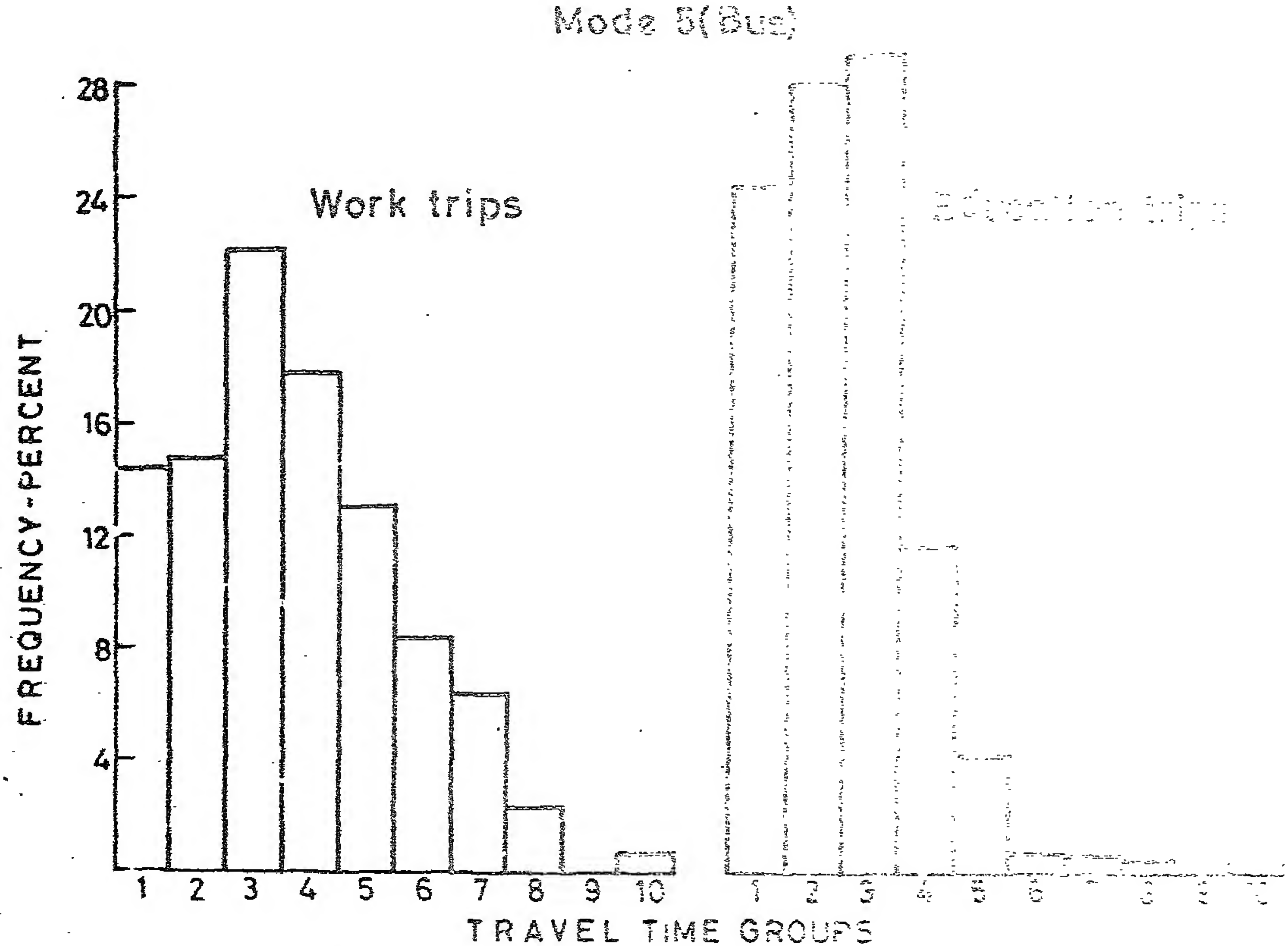


FIG-4·7 DECISION PROFILE OF TRAVEL TIMES FOR MODE 5(Bus)

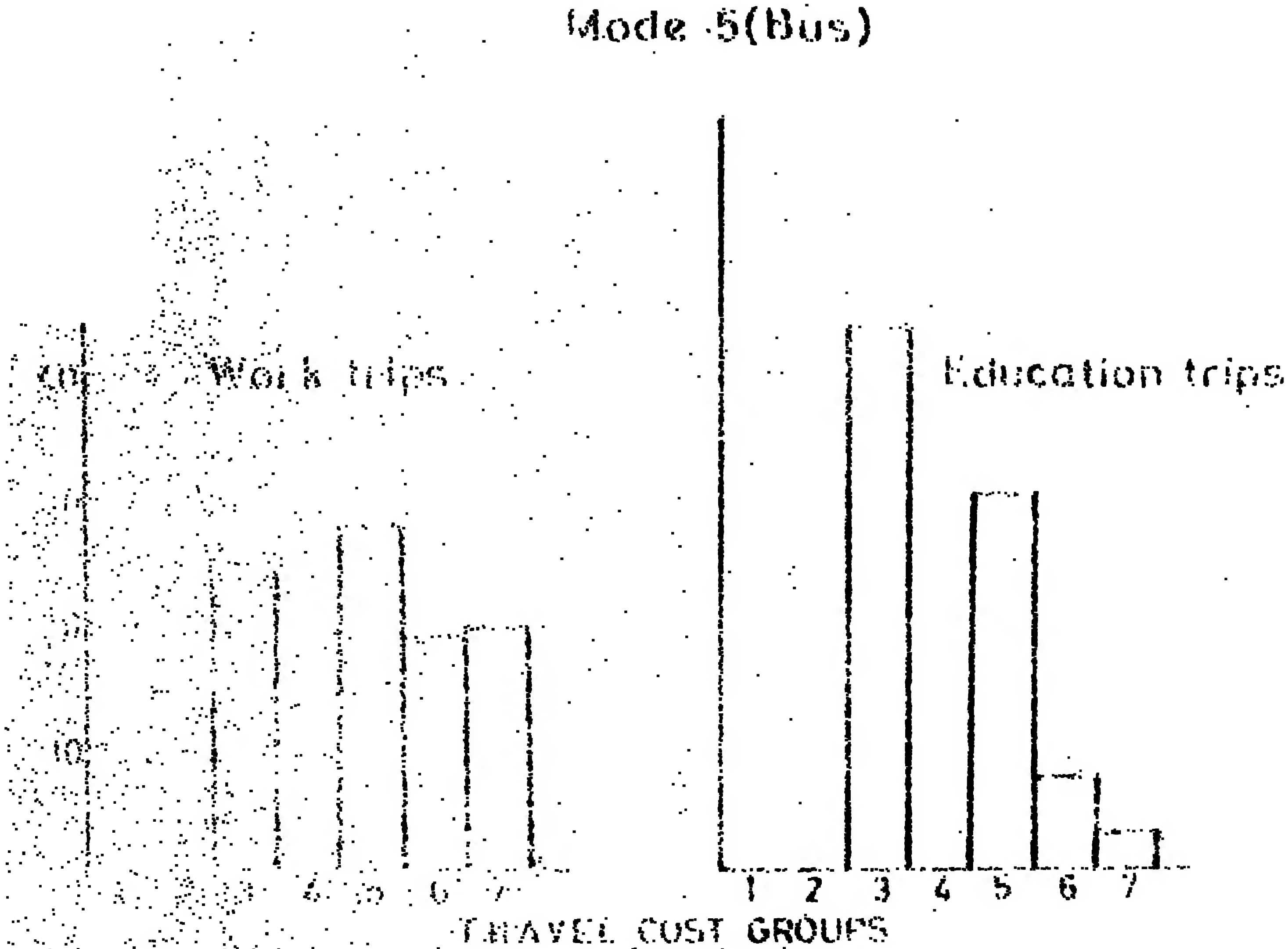
per the existing fare structure (minimum Rs. 0.50). The unit operating costs of the different modes considered in the study are:

Mode of travel	Cost 2 km. in Rupees
Bicycle	0.05
Motorcycle/scooter	0.42
Car	1.05
Cycle rikshaw	0.50

A samll dummy value of the travel cost by walk is incorporated. These costs are taken on the basis of the existing fuel cost and other maintenance charges. The travel costs are stratified into seven groups with different intervals (<0.25, 0.25-0.50, 0.50-0.75, 0.75-1.00, 1.00-1.50, 1.50-2.00, and > 2.00). For a particular status group the freqnency distributions of each of the travel cost group are represented in Fig. 4.8 for work and education trips. Similar decision profiles are derived for each of the status group by purpose. Here also the profiles are aggregated for all modes of travel and stratification is only by status group type. As the travel costs of the modes available for Kanpur city vary significantly the number of decision profiles for each of the modes are also derived and are shown in Fig. 4.9 for a particular mode.

4.5 MODEL ATTRIBUTES

The disaggregate model is a model of an individual decision making. The attributes of the various types of objects (individual, status group, zone and cluster of zones



IG.4 DECISION PROFILE OF TRAVEL COSTS FOR MODE 5(Bus)

between cordons) as considered in the model are given below:

(i) Individual (I)

- zone of origin/location of residence(i)
- income/education level (INC/EDL)
- status group (s)
- mode of travel (m)
- travel time group (tg)
- travel cost group (tc)
- destination group (d)

(ii) Status group(s)

- number of individuals in each status group ($NSOBS_s$)
- number of trips by different modes of travel($NTMOBS_{s,m}$)
- number of trips by different travel time groups($NTTOBS_{s,tg}$)
- number of employment/education opportunities for different destination groups ($NEOOBS_{s,d}$)
- number of trips by different travel cost groups($NTCOBS_{s,tc}$)

(iii) Zone (i)

- number of trips produced (NT_i)
- number of employment/education opportunities($NATOBS_i$)

(iv) Cluster of zones between cordons (c)

- income/education levels in cluster of zones between cordons ($INCOBS_c$).

4.6 MODEL STRUCTURE

4.6.1 Overview

The model simulates the modal split and trip distribution as an interconnected decision process at the individual level. It interprets the relations in the decision profiles as the probabilities and activates them with the use of random number generations. The components of the model as described in Fig. 4.10 describe the system of decision process with respect to individual work/education trips. The decision profiles which act as the major constraints on travel behaviour include:

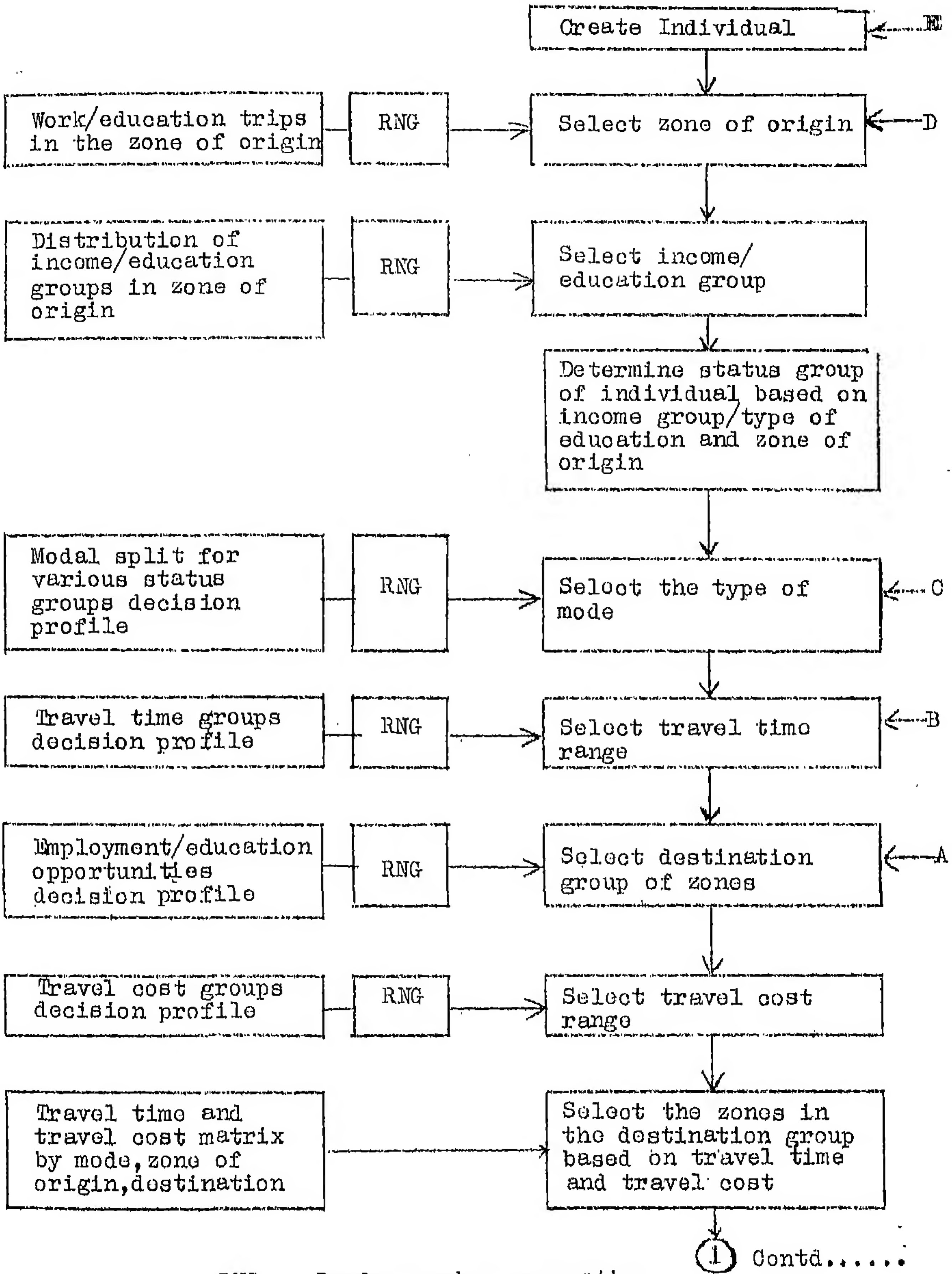
(i) distributions of productions and attractions in various zones,

(ii) distributions of income levels in the cluster of zones between cordons,

(iii) distribution for each status group of travel mode, travel time, travel cost, and employment/education opportunities, and

(iv) distributions of travel time and travel cost for each mode of travel.

The components of the model are activated in a specific order and with feedbacks as shown in Fig.4.10. The model is so structured that the order of activities of different components can be easily changed if required as per the framework of the study. The mode of transport is first selected in the proposed model and remains fixed within the decision process of the individual. If the object of study is to test the effect of improving the service level of certain public transport modes, then the decision on trip mode may be taken after selection of the decision cell and the possible feedbacks incorporated.



RNG = Random number generation

Fig. 4.10 contd...

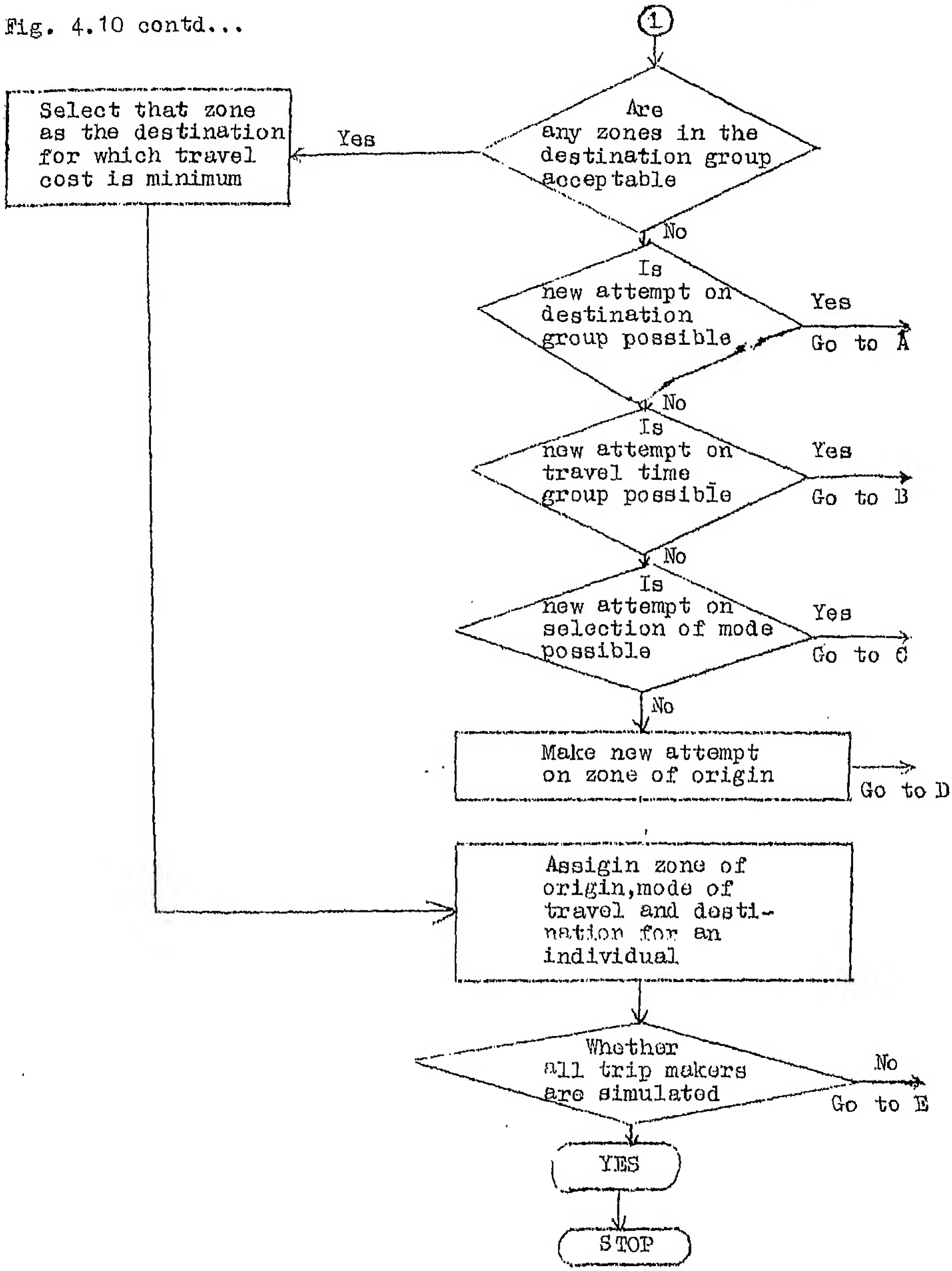


FIG. 4.10 : STRUCTURE OF THE MODEL

The simulation procedure first creates an individual. The zone of origin of the individual is selected from the decision profile of the trip productions in different zones and also with regard to the number of individuals already simulated in different zones of origin. Knowing the origin of the individual, the cluster of zones between cordons to which this origin belongs is identified. Individual is assigned an income level selected randomly from the decision profile of income level for the cluster of zone. From the origin zone and income level, the status group of the individual is identified for the work trips. When education trips are being simulated the level of education for the individual is selected from the relevant decision profile in the same manner as is done for selecting the income level. The status group of an individual performing an education trip is then identified as per the procedure given in Fig. 4.1.

The mode of travel for the individual is determined probabilistically by referring to the corresponding decision profile of his/her status group. As explained above,this selected mode of travel remains fixed during the decision cycle of individual's destination choice. If it so happens that the destination could not be assigned, then a new attempt on the selection of mode is made as per the procedure.

The procedure for choosing the destination is as follows. A travel time group is selected randomly from the decision profile of the status group for individual. This travel time group indicates a certain range of travel times . Using the

decision profile of employment/education opportunities, a set of destination zones that have the potential to attract the trip are identified. As there are a number of zones in the set, only those zones that can be reached within the selected travel time ranges are identified. When more than one destination zone are identified in the above procedure, a travel cost range is randomly selected from the decision profile for the concerned status group. Out of the already identified set of destinations, only those which can be reached within the selected travel cost range are identified and the one with minimum travel cost is selected provided it has already not been assigned the desired trip ends out of the previously simulated individuals. Otherwise the zone with the next higher travel cost within the selected travel cost range is then tested for selection.

The above procedure, thus, considers employment opportunity, travel time, and travel cost decision profiles as an integrated process. If no zone is selected in the above procedure then a new attempt is made in selecting the set of destination zones from the decision profile of employment/ education opportunities. A maximum number of attempts that can be made in the iteration may be specified in the model. If even after the maximum number of specified attempts on selection of destination groups, the final selection of destination could not be made, then a new attempt on the travel time group is attempted subject again to a maximum

number. Even if the above procedure does not select an acceptable destination then a new attempt is made on the selection of mode. If a certain maximum specified number of attempts on the mode does not result in final selection, then a new attempt can be made on the choice of origin of the individual.

The decision process of the simulation model is highly interconnected involving various decision profiles. The components of the system as used in the model are: (i) an individual, (ii) status group, (iii) zone, and (iv) cluster of zones between cordons. The attributes of these components are described in Section 4.5. The model is capable of simulting all the work/education trips of an urban area. For a large size metropolitan area it is economical to simulate only a sample of 5 percent.

The procedure of the model starts first by selecting an individual out of the population. The individual is simulated to determine the zone of origin, mode of travel and destination zone as an interconnected decision process. Next individual is selected only when the process for the previous one is over. The experiment stops when a certain fixed sample of individuals drawn from population is simulated. The structure of the model is such that once the decision process of the individual is complete, the individual is terminated and there is no further feedback to revise the decisions of an already simulated individual.

The model consists of the following component submodels:

(i) Status group submodel

(ii) Mode choice submodel

(iii) Distribution submodel.

These submodels are shown in the Fig. 4.11 and explained in the following sub-sections.

4.6.2 Status Group Submodel

This submodel identifies the status group by purpose of the individual being simulated. The submodel involves firstly the selection of origin and his/her household income/education level. When an individual I to be simulated is created, the zone of origin for the individual is drawn randomly from the decision profile of trip productions for different zones. Let the zone of origin drawn be $i = Zone_I$. This zone of origin is accepted if the number of trip productions already simulated from a zone do not exceed the total number of trip productions by purpose of that zone, that is, if,

$NTSIM_i < NT_i$, then select i as the zone of origin

where $NTSIM_i$ = number of trips produced in zone i out of the already simulated trips

NT_i = total number of trips produced in zone i by purpose.

If the zone of origin i does not satisfy the above condition, then another zone of origin is drawn from the decision profile till final selection of i is made. The trip productions by purpose of the simulated trips for the zone i is now updated, i.e.,

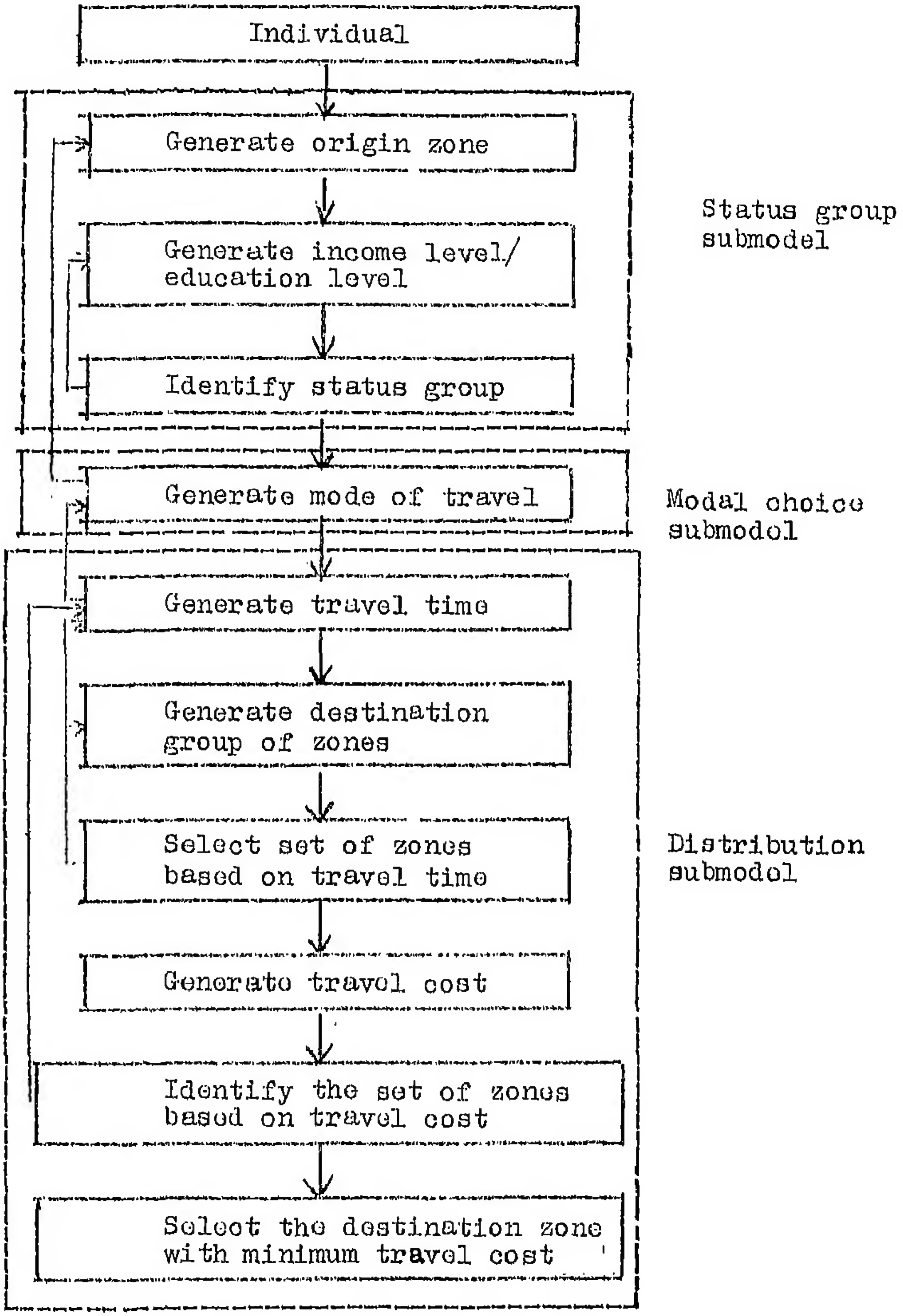


FIG. 4.11 : INTERACTION OF SUBMODELS

$$NTSIM_i = NTSIM_i + 1 \qquad .. (4.1)$$

The income level of the individual I is drawn from the decision profile of the cluster of zones in which the individual resides. Now let c be the cluster of zones that includes i ($c = CLUSTER_i$). This cluster is based on the location of zone with regard to the cordons (c = 1 for zones within CBD , c = 2 for zones between CBD and intermediate cordon; and c = 3 for zones between intermediate and outer (cordons). From the decision profile of the income level for the cluster of zones c the income level of the individual, INC_I, is generated. In the same manner the education level of the individual, EDL_I, is generated from the relevant decision profile. The zone of origin and income/education level identifies the status group of the individual as shown in Fig. 4.2. Mathematically this status group s of the individual is expressed as

$$s = (INC_I - 1) * 3 + c \text{ for work trips} \qquad .. (4.2)$$

$$s = (EDL_I - 1) * 3 + c \text{ for education trips} .. (4.3)$$

Like the zonal trip productions it may be tested that the number of individuals simulated in a particular status group do not exceed the observed number for that group , i.e.,

$$NSSIM_s < NSOBS_s$$

where, $NSSIM_s$ = number of individuals simulated in status group s out of the already simulated individuals

$NSOBS_s$ = total number of individuals in status group s by purpose

The inclusion of the above status group constraint is optional. Once status group is accepted, the number of individuals simulated in a status group are updated, i.e.,

$$NSSIM_s = NSSIM_s + 1 \qquad .. (4.4)$$

4.6.3 Mode Choice Submodel

This submodel selects the mode of travel for the status group of the individual as identified in sub-section 4.6.2. The mode of travel m for the status group is drawn randomly from the decision profile of the mode of travel. Once this mode is finally accepted, the number of simulated individuals using the mode m is updated, i.e.,

$$NTMSIM_{s,m} = NTMSIM_{s,m} + 1 \qquad .. (4.5)$$

where, $NTMSIM_{s,m}$ = number of trips by mode m for status group s out of the already simulated trips.

4.6.4 Distribution Submodel

This submodel identifies the final destination zone of the individual being simulated. It involves, firstly, the selection of travel time group. The travel time group, tg, is selected randomly from the decision profile of the travel time for the relevant status group. The travel time for the set of destination zones to be selected should be within this travel time range of (tg-1) and tg. The destination group of zones is identified from the decision profile of the employment/education opportunities. From the number of zones n of the destination group d , a set of zones that have the potential to attract the trips are identified, i.e.,

$$\text{NATSIM}\ (\text{ZONES}_{d,n}) < \text{NAT}\ (\text{ZONES}_{d,n})$$

where, $\text{NATSIM}(\text{ZONES}_{d,n})$ = number of trip attractions in the nth zone of the destination group d out of the already simulated trips,

$\text{NAT}(\text{ZONES}_{d,n})$ = total number of trip attractions in the nth zone of the destination group d.

If the above condition is not satisfied, then another destination group of zones is drawn from the decision profile of the employment/education opportunities. For the zones which are identified from the travel time by the selected mode m from zone of origin i to different zones in the destination group are tested with the travel time ranges, i.e.,

$$\text{TIM}_{tg-1} < \text{TT}(m,i,\text{ZONES}_{d,n}) \leq \text{TIM}_{tg}$$

where, TIM_{tg-1} = travel time for the (tg-1)th group,

$\text{TT}(m,i,\text{ZONES}_{d,n})$ = travel time for nth zone in the destination group d from origin i by mode m,

TIM_{tg} = travel time for the tg th group.

If the above condition is not satisfied for any of the identified destinations , it is decided to select another destination group from the decision profile. If only one zone is accepted, that is taken as the destination zone j. When more than one destination zone are identified in the above procedure, a travel cost is randomly selected from the decision profile of the concerned status group. The set of zones which can be reached by the selected travel cost range are identified, i.e. ,

$$\text{TCOS}_{tc-1} < \text{TC}\ (m,i,\text{ZONES}_{d,n}) < \text{TCOS}_{tc}$$

where, $TCOS_{tc-1}$ = travel cost for the (tc-1)th group,

$TC(m,i,ZONES_{d,n})$ = travel cost for nth zone in the destination group d from zone of origin i by mode m,

$TCOS_{tc}$ = travel cost for the tc th group.

Out of the destination zones n in the group d the destination zone j with the minimum travel cost is selected if,

$$NATSIM_j < NAT_j$$

where, $NATSIM_j$ = number of trip attractions to zone j out of the already simulated trips,

NAT_j = total number of trip attractions to zone j by purpose.

The trip attractions by purpose of the simulated trips to the destination zone j are now updated.

$$\text{i.e., } NATSIM_j = NATSIM_j+1 \qquad .. (4.6)$$

In the above procedure, if the destination zone j is not selected, new attempt on the destination group of zones is made subject to a maximum of five attempts. If the procedure still fails to select the destination zone, new selections are made in order for travel time, mode of travel, and zone of origin. A maximum of five attempts are made in each selection.

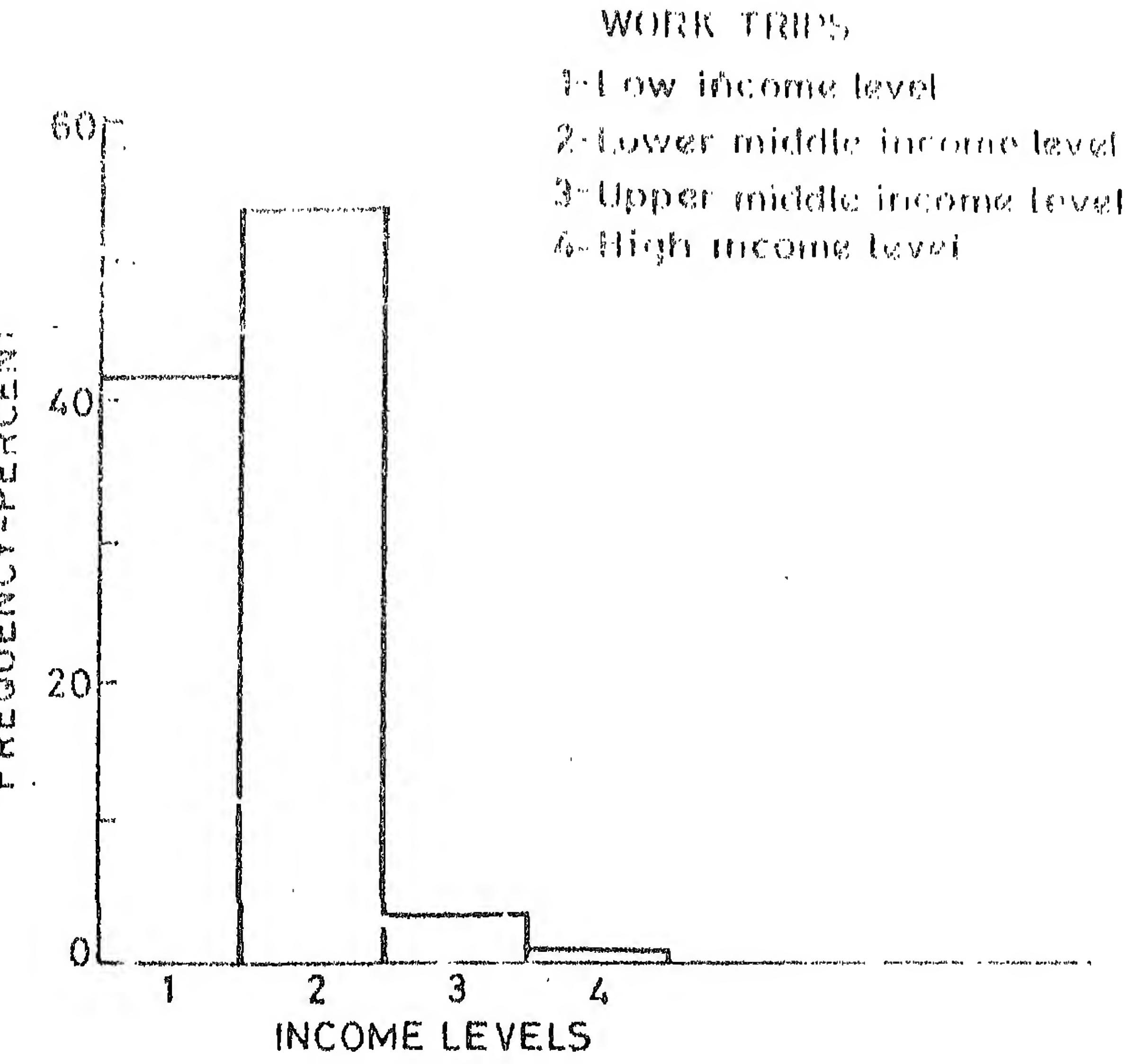
4.7 MODEL CALIBRATION

The submodels which are described in sub-sections 4.6.2 to 4.6.4 require calibration against the observed data. The household data which is required for these models are collected as shown in Section 2.3. It may also be necessary to calibrate the values of some parameters and decision

thresholds in the light of the model results during validation. The decision profiles for the dependent variables like mode of travel, employment/education opportunities, travel time, and travel cost based on status group along with the decision profiles for travel time and travel cost based on mode of travel are calibrated in sub-section 4.4.3. The distributions of number of individuals in a particular status group affect the system. Further as the status group depends upon the income/education level and the residential location, it is necessary to have the decision profiles for income/education levels.

Income/Education level: The decision profile for the income levels in cluster of zones between cordons is derived from the household survey data. Household income in classified into four levels (low income group, lower middle income group, upper middle income group, high income group) and so is the case for education levels (primary, secondary, degree, vocational). A representative decision profile of income level distributions for the cluster of zones within CBD is shown in Fig. 4.12. This profile gives the frequency in percentage of individuals by each of the income level.

Individuals by status group: The decision profiles for the number of individuals in different status group are derived from household income/education level and location of residence and a representative one is shown in Fig. 4.13. The decision profiles for the other dependent variables are already explained in sub-section 4.4.3.



FIG·4·12 DECISION PROFILE FOR INCOME LEVELS WITHIN CENTRAL CORDON

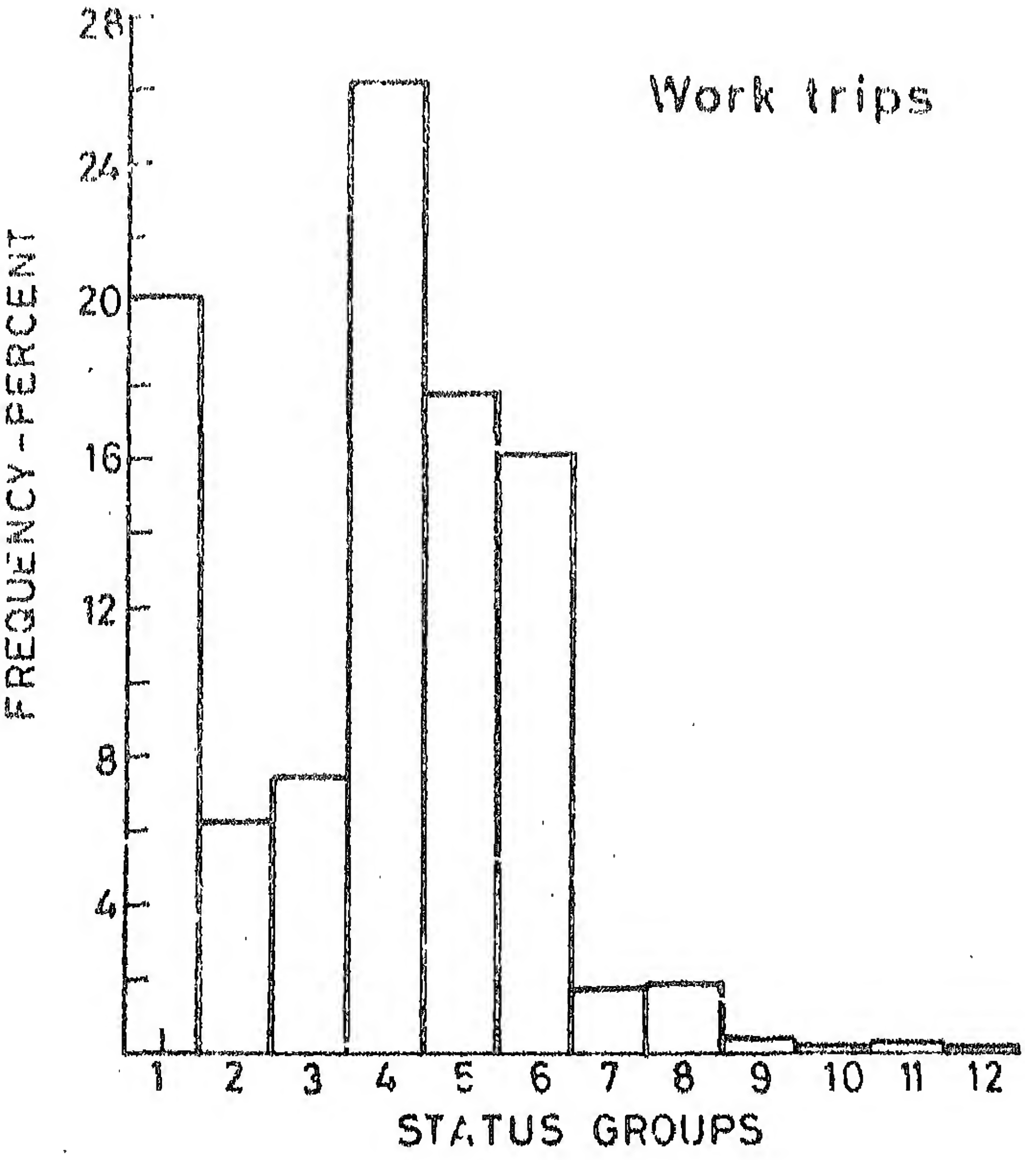


FIG.4.13 DECISION PROFILE OF STATUS GROUPS

4.8 MODEL VALIDATION

4.8.1 General

The model can be used to simulate future conditions only after it has been validated and found to be a reasonable representation of the actual real system under study (Bulgren, 1982). Validation involves the verification that the structure of the model is correct and parameter estimates are reasonable. Sometimes validation may also involve parameter estimation. Generally the validation of the model is based on available information and an agreement between the output of the model with earlier observations. This process of validation is a straight forward one but may present some difficulties in carrying out. Since the data are usually probabilistic, for small lengths of simulation runs, the variability of the model output would be large. In this case one has to simulate the model several times with different sequence of random numbers and obtain the range of variation among these. If the outputs to be compared are sample means one can use number of statistical tests available to measure the discrepancy.

In this study it is decided to validate the model by comparing the model outputs with the observed household survey data. The model consists of the following submodels:

(i) Status group submodel

(ii) Mode choice submodel

(iii) Distribution submodel

These models that are based on available information must be validated seperately. Using the overall simulation results it

is necessary to calibrate the values of some of the parameters. The overall simulation model is a test of how well well the submodels have been assembled into a realistic structure of the system.

4.8.2 Strategy for Validation

The model simulates the decision process of the individual and selects the travel mode and trip end. The decision process is based upon the status group of the individual. For proper validation of the system it is necessary to see that each of the submodels,namely,status group submodel, modal split submodel, and distribution submodel are validated seperately so as to ensure that they have been assembled into a realistic structure of the system. Properties to be used as measures of effectiveness in validation should be such that they represent the output of the whole system and can also be observed from the real world data. These properties should be such that significant disagreement between observed and simulated values implies that the model is unsatisfactory and an insignificant disagreement implies satisfactory nature of the model. The properties selected for validation of the various submodels in this study are:

Status group submodel

(i) income/education level distribution for the cluster of zones between various cordons of the study,

(ii) trip productions for the cluster of zones,

(iii) trip productions by each status group.

Status group is based upon the residential location and income/education level of the individual. Properties (i) and (ii) are a subset of the status group submodel whereas property (iii) represents the characteristics of the different status groups.

Modal split submodel

(i) trips by purpose by different modes of travel,

(ii) trips performed by different modes for each of the status groups.

Property (i) is validation of the aggregated value of the results from modal split model model whereas property (ii) represents the stratification of modal trips into each of the status groups. These properties include the interaction of status group and modal split models also.

Distribution submodel

An ideal validation of the trip distribution submodel can be by comparing the simulated trips for different O-D pairs with the observed ones. This comparison for each element of the O-D matrix is at a highly microlevel. Besides this, the other properties adopted for empirical validation are the frequency distributions of trips for different employment/education opportunities, travel times and travel costs.

The model is used to simulate 12485 individuals performing work trips and 8952 individuals making education trips in the city of Kanpur. These numbers correspond to the 5 percent sample size as recorded in the household survey data.

The aim of these simulation experiments is to have an empirical validation of the model by comparing the simulation results with the observed data. The validation is quite extensive involving properties related to status groups, mode choice and trip distribution submodels and their interconnections. The following sections show in detail, the comparisons of the simulated and observed data for various properties.

4.8.3 Status Group Submodel

The status group of the individuals depends upon the residential location and income/education level. The aggregate measure for validation of the submodel is the comparison between the observed and simulated trip productions in each status group. Besides this the income/education level distributions in the cluster of zones between cordons are also compared.

The distribution of observed and simulated income levels for each cluster of zones between the cordons is given in Table 4.1. The values in an income level group represent the percentage of the individuals belonging to that income level of the cluster. The distributions of observed and simulated values of educational levels for the different clusters are given in Table 4.2. The cumulative frequencies of income levels and educational levels for the first cluster of zones within the CBD are also shown in Fig. 4.14. The results indicate that the distributions of income level and educational levels of the simulated individuals are highly close to the observed ones for each set of zones. This indicates that the

TABLE 4.1 : DISTRIBUTION OF INCOME LEVELS IN CLUSTER OF ZONES BETWEEN CORDONS

Cluster no.	Income level 1		Income level 2		Income level 3		Income level 4	
	Observed percent	Simulated percent	Observed percent	Simulated percent	Observed percent	Simulated percent	Observed percent	Simulated percent
Cluster 1 (within CBD)	41.67	41.75	53.79	53.74	3.52	3.47	1.02	1.04
Cluster 2 (between CBD and intermediate cordon)	23.57	24.06	66.99	68.59	6.87	6.69	2.56	0.65
Cluster 3 (between intermediate and outer cordons)	30.32	28.68	65.76	67.03	3.65	4.00	0.26	0.29

TABLE 4.2 : DISTRIBUTION OF EDUCATIONAL LEVELS IN CLUSTER OF ZONES BETWEEN CORDONS

Cluster no.	Education level 1		Education level 2		Education level 3		Education level 4	
	Observed percent	Simulated percent	Observed percent	Simulated percent	Observed percent	Simulated percent	Observed percent	Simulated percent
Cluster 1 (within CBD)	49.51	49.10	44.26	44.61	5.28	5.32	0.95	0.96
Cluster 2 (between CBD and intermediate cordon)	50.19	51.05	41.43	43.20	6.70	4.88	1.68	0.87
Cluster 3 (between intermediate and outer cordons)	59.29	57.90	34.73	35.82	5.32	5.59	0.66	0.69

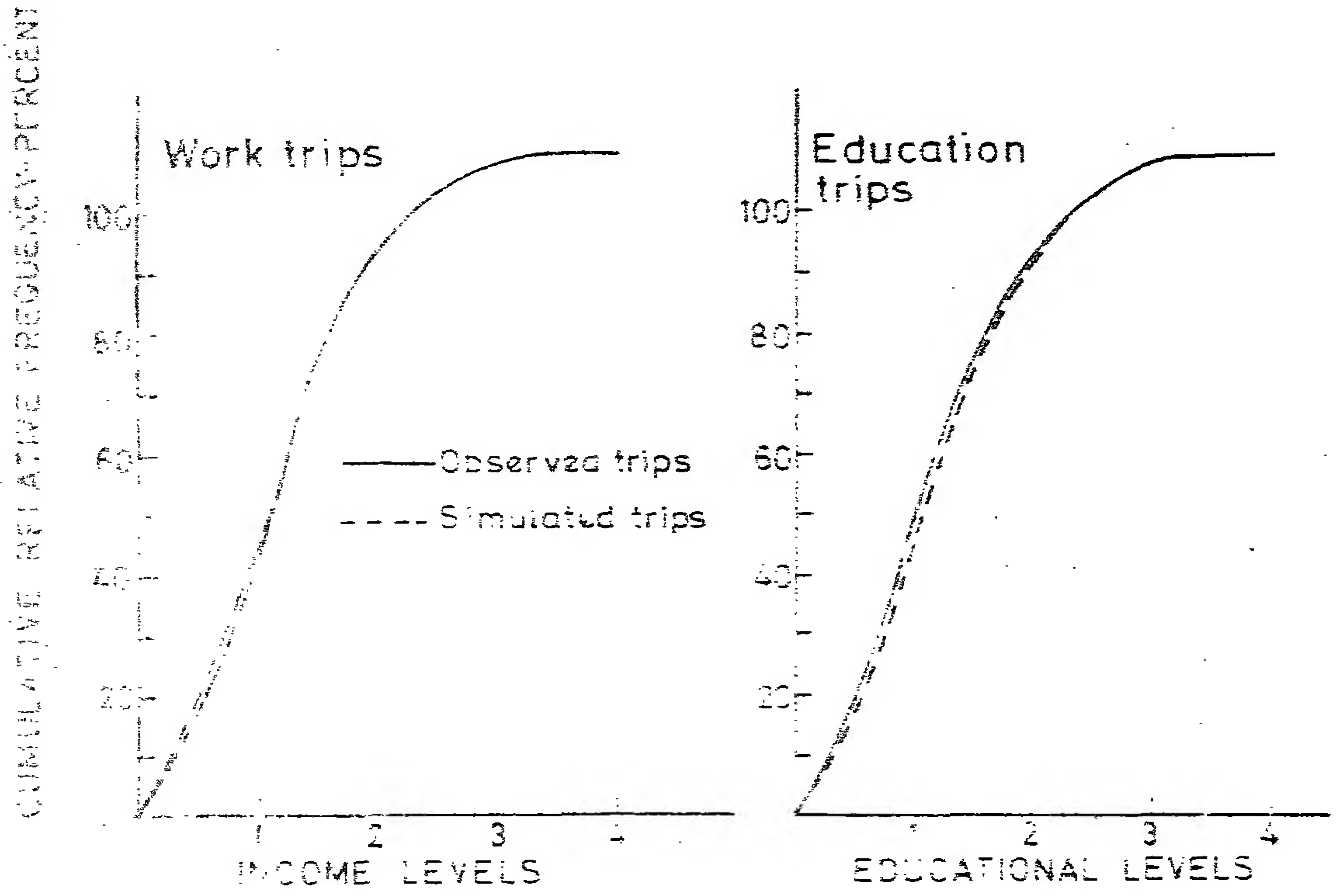


FIG 4.14 CUMULATIVE FREQUENCY DISTRIBUTIONS OF INCOME/ EDUCATIONAL LEVELS WITHIN CORDON 1

sample size chosen for simulation is sufficient enough as the individuals drawn by simulation have similar income/education characteristics as the observed ones.

The simulation algorithm while identifying the zone of origin of an individual tries to maintain the production constraints of that zone. To test whether the simulated productions in each of the cluster of zones compare with the observed ones or not, the simulated and observed trip productions for each of the clusters are given in Table 4.3. The comparison indicates that the two are highly close to one another. The mean percentage difference for both work trips and education trips is about -0.5 , the maximum value being 4.96 for work trip productions in cluster number 3. These comparisons indicate that the simulated trip productions aggregated to the cluster level compare extremely well with the observed ones.

The status group submodel decides the status group of the individuals. The distributions of trip productions for each of the status groups, both for work and education trips are derived from the simulated results and they are shown in Figs. 4.15 and 4.16 alongwith the observed values. These distributions are represented in proportion of the total trips. The simulated and observed proportions are quite close to one another. The analysis of these observed and simulated trip productions for each of the status groups are given in Table 4.4 and comparison is also shown in Fig. 4.17. The results indicate that the variation between the observed and simulated values is within $\pm$ 5 percent in general. It is only in a couple of cases that

TABLE 4.3 : COMPARISON OF TRIPS IN CLUSTER OF ZONES BETWEEN CORDONS

Cluster no.	Work trips		Percentage difference	Education trips		Percentage difference
	Observed trips(O)	Simulated trips (S)	$(\frac{S-O}{O} \times 100)$	Observed trips(O)	Simulated trips(S)	$(\frac{S-O}{O} \times 100)$
Cluster 1 (within CBD)	6103	6200	+ 1.59	4112	4191	+ 1.92
Cluster 2 (between CBD and intermediate cordon)	3318	3373	+ 1.66	2865	2827	- 1.33
Cluster 3 (between intermediate and outer cordons)	3064	2912	- 4.96	1975	1934	- 2.08
Mean percentage differende		=	- 0.57			- 0.50

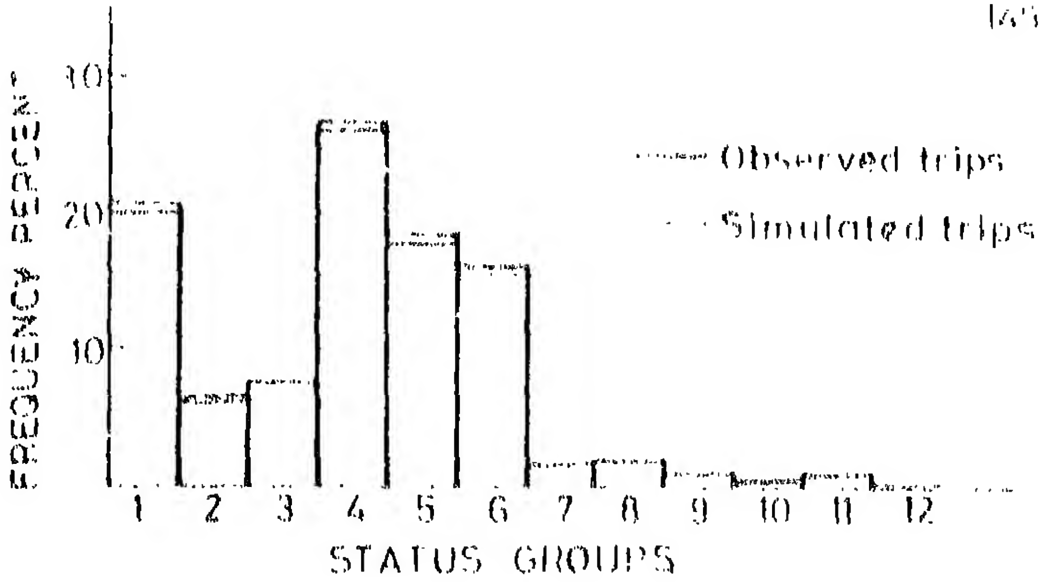
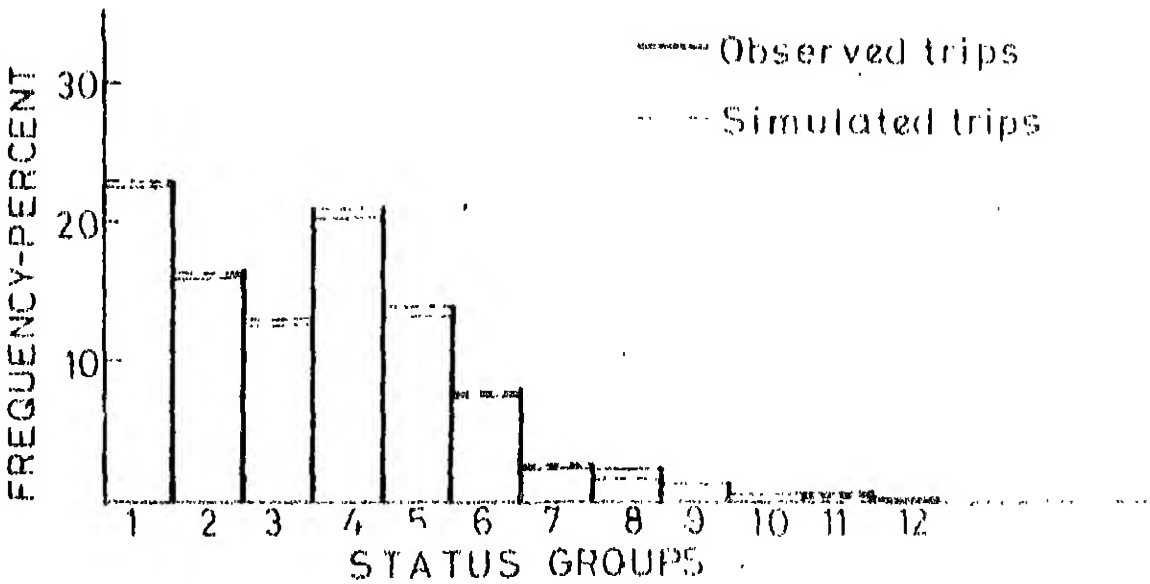


FIG 4·15 FREQUENCY DISTRIBUTION OF STATUS GROUPS FOR WORK TRIPS



FIG· 4·16 FREQUENCY DISTRIBUTION OF STATUS GROUPS FOR EDUCATION TRIPS

TABLE 4.4 : DISTRIBUTION OF TRIP PRODUCTIONS FOR EACH STATUS GROUP

Status gruop no.	Work trips: Observed trips(O)	Work trips: Simulated trips(S)	Percentage difference $(\frac{S-O}{O} x100)$	Education trips: Observed trips(O)	Education trips: Simulated trips(S)	Percentage difference $(\frac{S-O}{O} x\ 100)$
1	2543	2588	+ 1.77	2036	2058	+ 1.08
2	782	811	+ 3.71	1438	1444	+ 0.42
3	929	835	-10.12	1171	1120	- 4.36
4	3283	3332	+ 1.49	1820	1870	+ 2.75
5	2223	2314	+ 4.09	1187	1220	+ 2.78
6	2015	1953	- 3.07	686	693	+ 1.02
7	215	216	+ 0.47	217	223	+ 2.76
8	228	225	- 1.32	192	138	-28.13
9	112	116	+ 3.57	105	108	+ 2.86
10	62	64	+ 3.22	39	40	+ 2.50
11	85	23	-72.94	48	25	+47.92
12	8	8	0.00	13	13	0.00
Mean percentage difference		=	- 5.77			- 5.35
Mean absolute percentage difference		=	8.80			8.05

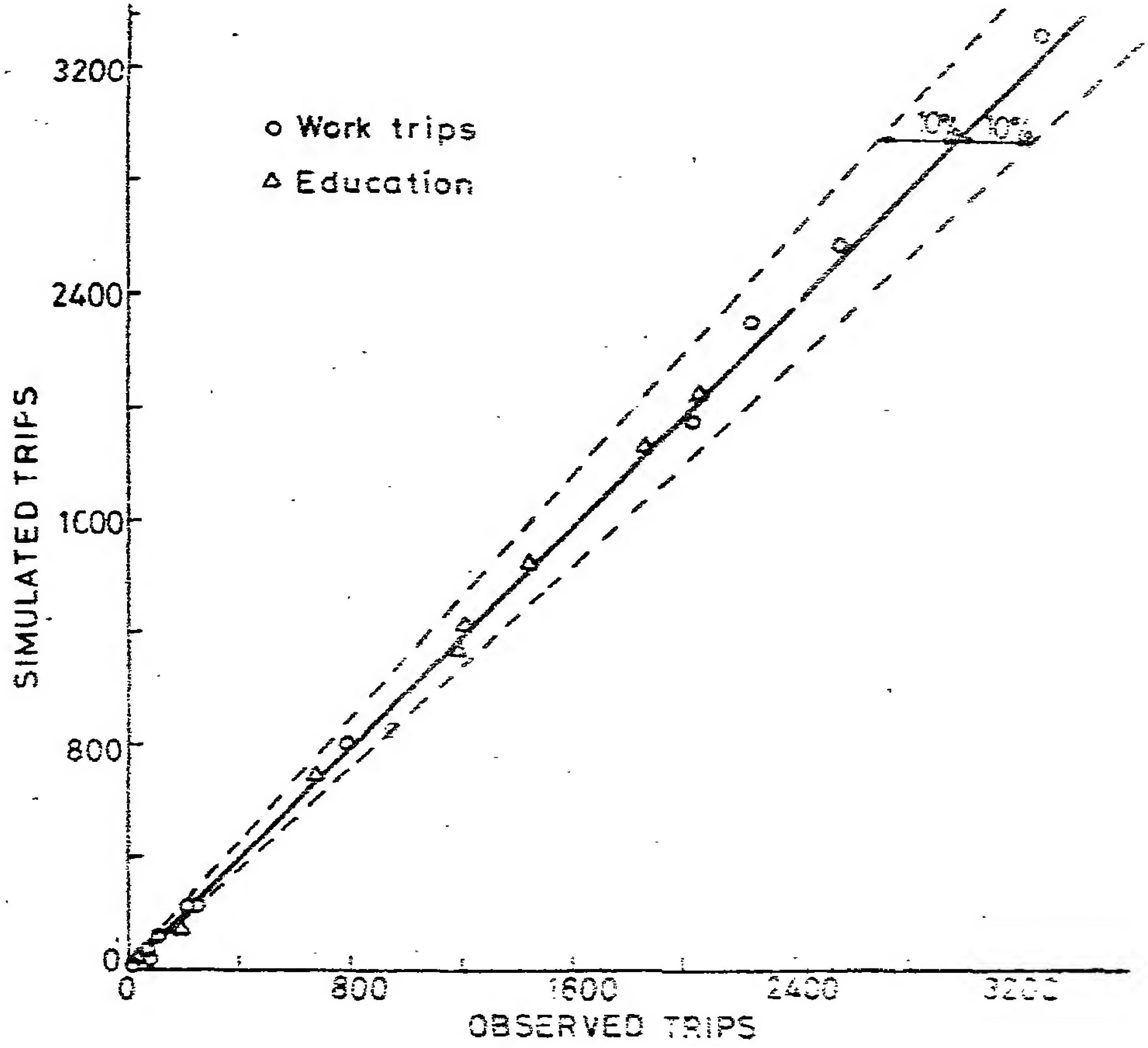


FIG-4.17 COMPARISON OF OBSERVED AND SIMULATED TRIPS BY DIFFERENT STATUS GROUPS

the variation is more than 10 percent. But this happens only when the sample size is quite low. Further the Fig.4.17 indicates that the data points lie on both sides of the 45 degree line indicating that there is no general bias in the simulation algorithm. The results indicate that the status group submodel simulates the real world situation sufficiently well and the simulated size of individuals that form 5 percent of the population is significant enough for the case of Kanpur city. By increasing the sample size the simulated results may further come close to the observed ones. This will however involve increased computation time without any appreciable gain in the overall outcome.

4.8.4 Mode Choice Submodel

The decision of mode choice of the individual being simulated is governed by the procedure discussed in sub-section 4.6.3. The number of trips performed by each mode are determined and are compared with the observed trips so as to have an empirical validation of the mode choice submodel. Figs.4.18 and 4.19 show the cumulative relative frequencies of the trips made by different modes for work and education purposes. These figures indicate that the simulated frequencies are very close to the observed ones for all the modes. Analysis of these results given in Table 4.5 indicates that the mean percentage difference between the observed and the simulated trips of all modes is only -1.01 for work trips and +1.19 for education trips, while the corresponding mean absolute percent differences are 4.14 and 1.44 respectively. The variations are little higher

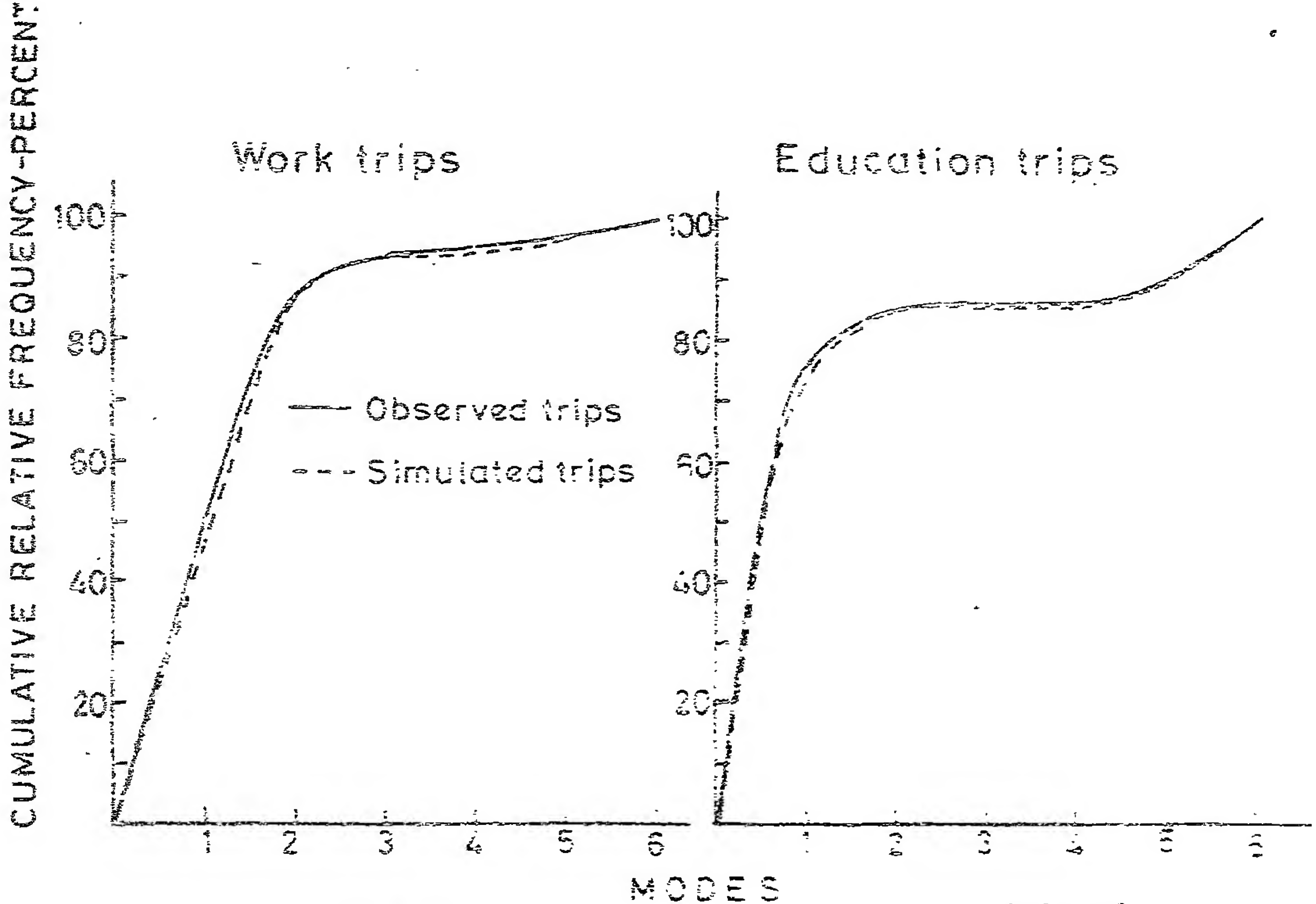


FIG. 4.18

FIG. 4.19

CUMULATIVE FREQUENCY DISTRIBUTION OF TRIPS BY DIFFERENT MODES

TABLE 4.5 : ANALYSIS OF TRIPS BY DIFFERENT MODES OF TRAVEL

Mode	Work trips		Percentage difference	Education trips		Percentage difference
	Observed trips(O)	Simulated trips(S)	$(\frac{S-O}{O} \times 100)$	Observed trips(O)	Simulated trips (S)	$(\frac{S-O}{O} \times 100)$
Walk	5953	5737	− 3.63	6844	6791	− 0.77
Bicycle	4967	5164	+ 3.97	812	833	+ 2.59
Motorcycle/ scooter	762	783	+ 2.76	31	31	0.00
Car	93	82	−11.83	13	13	0.00
Bus	298	303	+ 1.68	383	394	+ 2.87
Cycle rikshaw	412	416	+ 0.97	869	890	+ 2.42
Mean percentage difference =			− 1.01			+ 1.19
Mean absolute percentage difference=			4.14			1.44

only in those cases where the sample size is quite low. This shows that the model is capable of representing the real situation. The above comparison of mode choice is aggregated for all the trips in the region of study. As the mode choice submodel is interconnected to the status group submodel the simulated values of trips by different modes for each status group are compared with the observed values. Figs. 4.20 to 4.22 indicate comparison of the observed and simulated trips performed by bicycle, motorcycle/scooter and cycle rikshaw. Each point in the figure corresponds to a particular status group by purpose. These figures indicate that the simulated values are well within $\pm$ 10 percent of the observed ones. Further the points are spread out on both sides of the 45 degree line.

Tables 4.6 and 4.7 give the relative frequencies by different modes of travel for each of the status groups. For work trips the mean percentage differences are within $\pm$ 4 percent for all the status groups except for status group number 11. This particular status group corresponds to residential location of high income individuals in the area between CBD and intermediate cordon. The number of individuals belonging to this status group is very small. For education trips the mean percentage differences are even smaller than those for work trips. The above comparisons indicate that the mode choice submodel interconnected with the status group submodel gives sufficiently reasonable results showing the ability of the model to replicate the system.

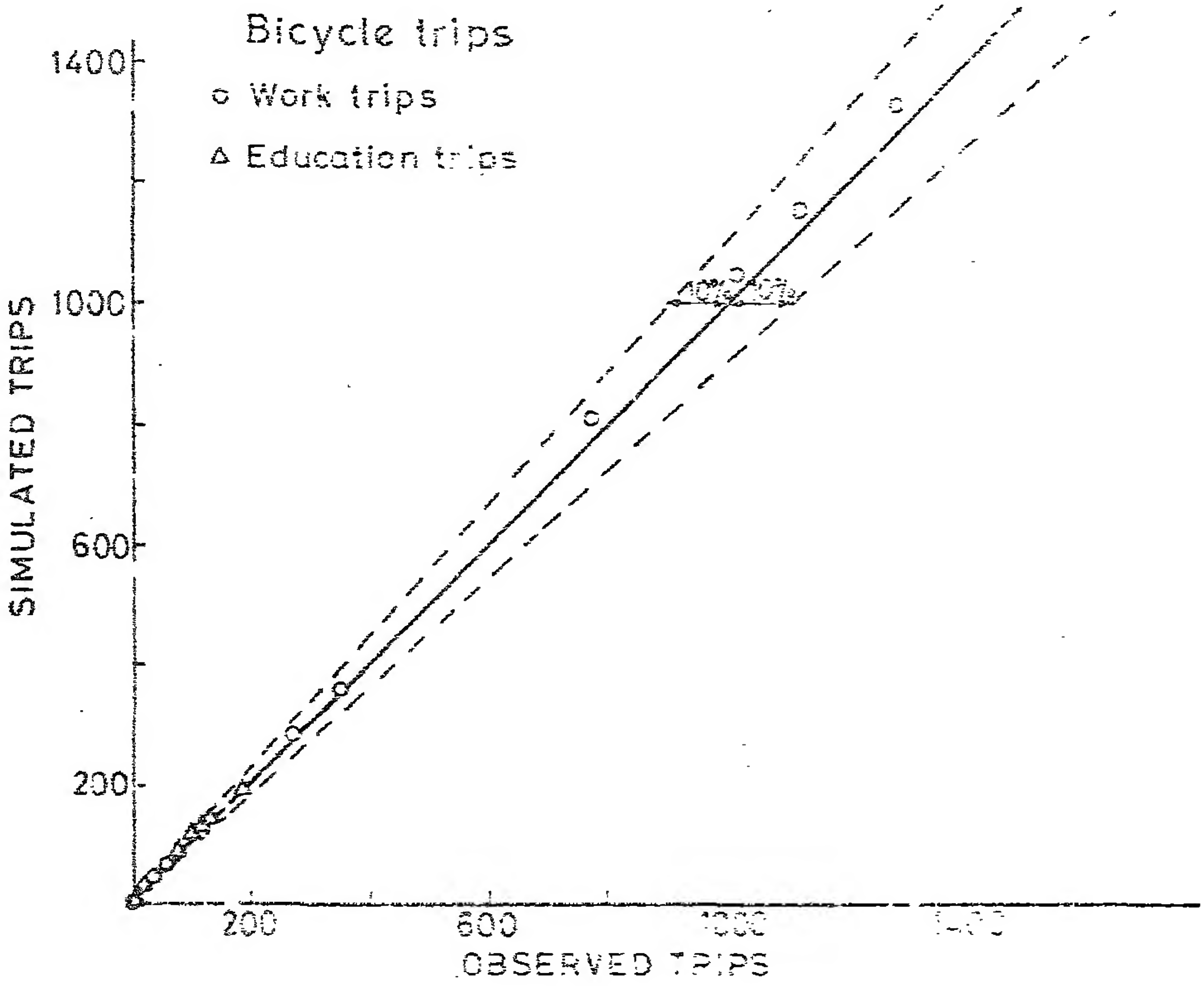


FIG.4.20 COMPARISON OF OBSERVED AND SIMULATED TRIPS BY DIFFERENT STATUS GROUPS

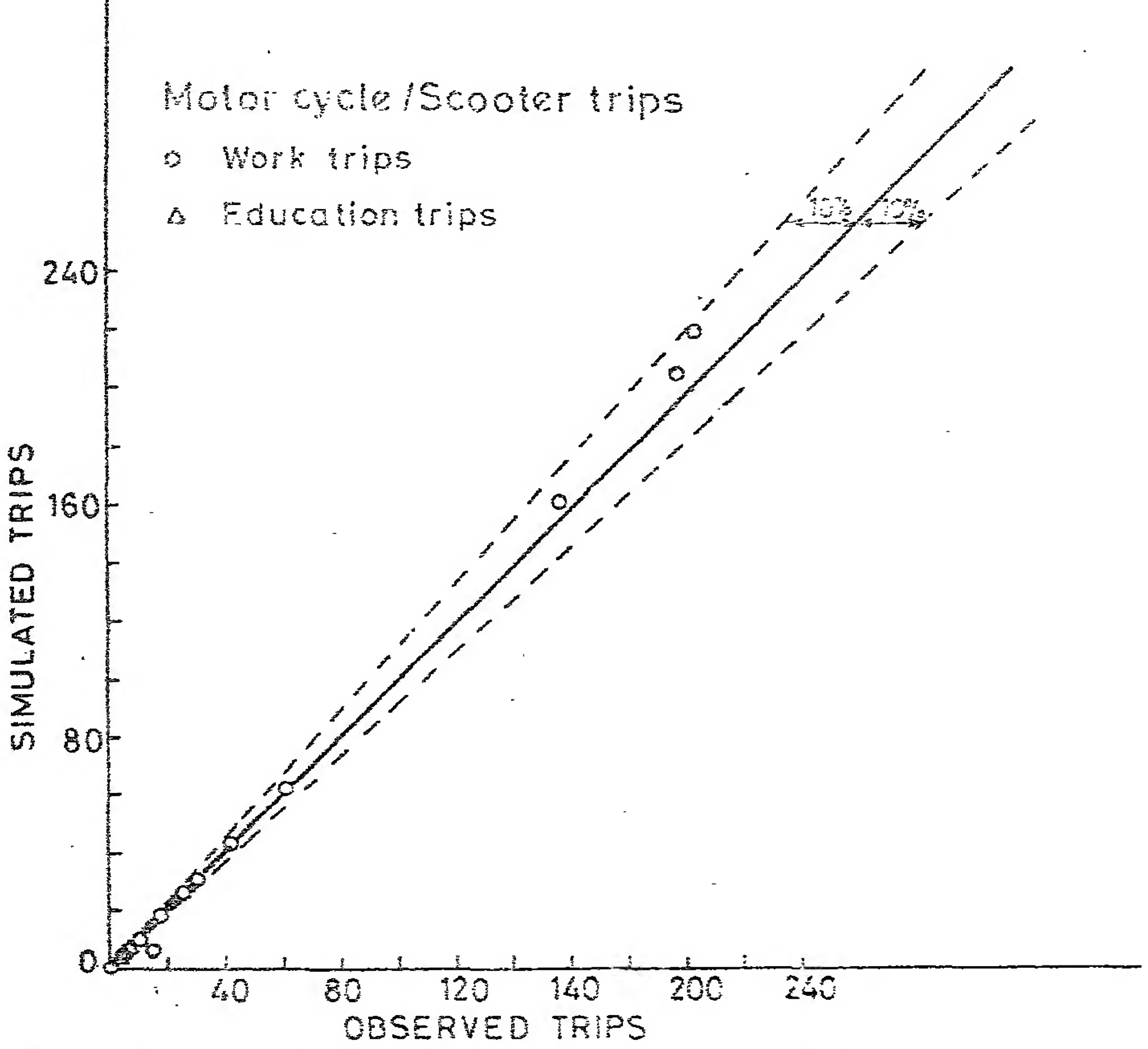


FIG. 4.21 COMPARISON OF OBSERVED AND SIMULATED TRIPS BY DIFFERENT STATUS GROUPS

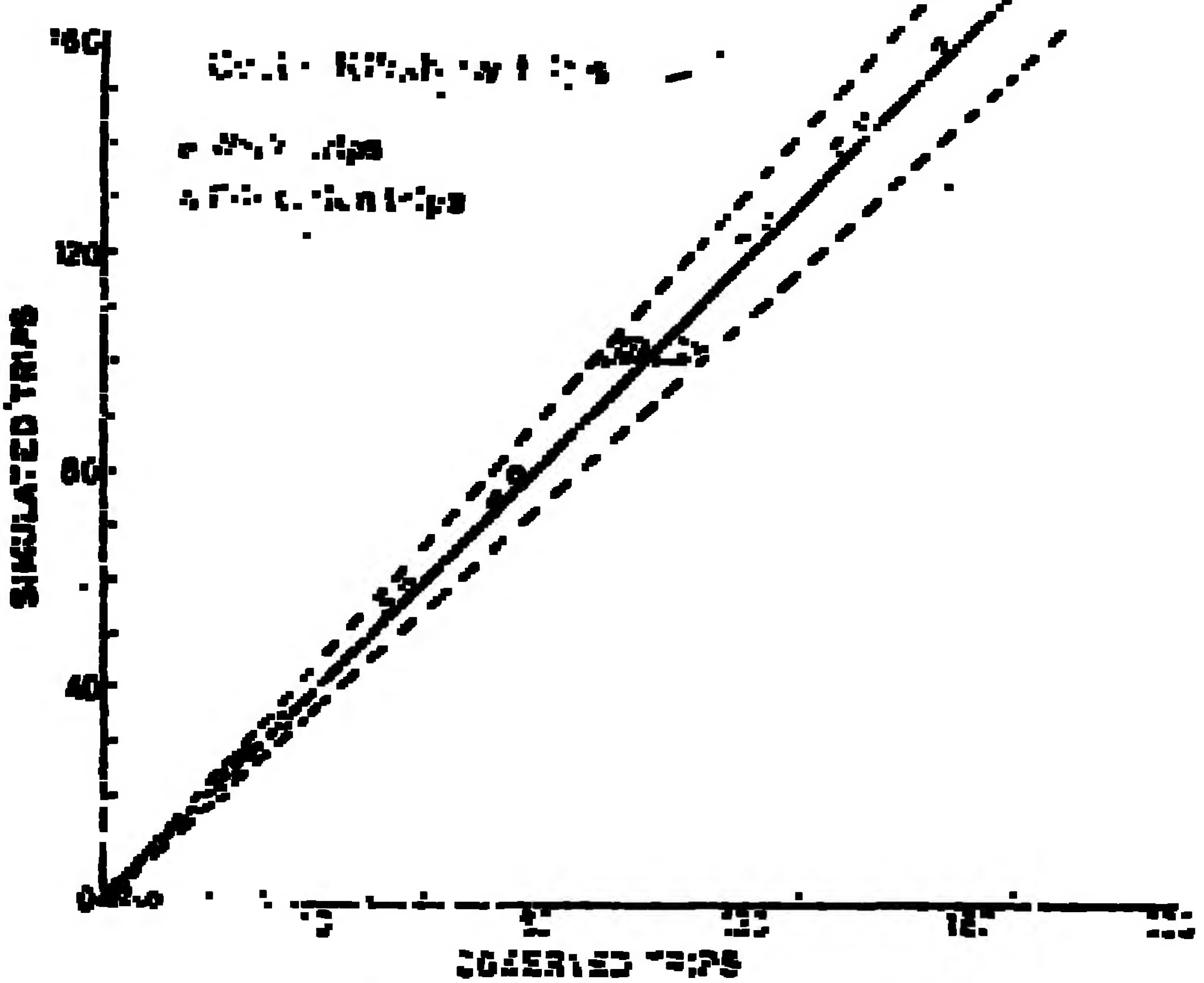


FIG-4.22 COMPARISON OF OBSERVED AND SIMULATED TRIPS
BY DIFFERENT STATUS GROUPS

TABLE 4.6 : WORK TRIP MODE CHOICE FOR EACH STATUS GROUP

Status group no.		Proportion of trips by different modes						Mean percentage difference	Mean absolute percentage difference
		Walk	Bicycle	Motorcycle/ scooter	Car	Bus	Cycle rikshaw		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	O*	64.53	30.79	1.02	0.00	0.63	3.03	+ 3.13	3.13
	S**	63.75	31.47	1.05	0.00	0.64	3.09		
2	O	50.51	44.63	0.13	0.13	1.79	2.81	+ 1.73	3.25
	S	50.64	44.74	0.13	0.13	1.79	2.56		
3	O	61.57	29.71	1.18	0.22	3.66	3.66	- 1.96	4.27
	S	55.67	34.37	1.37	0.45	4.23	4.11		
4	O	48.95	38.93	6.15	0.40	1.95	3.62	+ 2.48	2.48
	S	47.66	39.91	6.31	0.41	2.00	3.72		
5	O	36.80	50.11	7.02	0.45	3.24	2.38	+ 3.33	3.33
	S	36.81	50.11	7.00	0.43	3.24	2.38		
6	O	33.65	50.27	9.78	0.15	3.33	2.83	- 0.17	5.53
	S	28.77	53.97	10.50	0.16	3.57	3.04		

Contd....

TABLE 4.7 : EDUCATION TRIP MODE CHOICE FOR EACH STATUS GROUP

Status group no.		Proportion of trips by different modes						Mean percentage difference	Mean absolute percentage difference
		Walk	Bicycle	Motorcycle/ scooter	Car	Bus	Cycle rikshaw		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	O*	84.00	6.00	0.10	0.20	2.40	7.30	+ 1.43	1.43
	S**	83.20	6.25	0.15	0.20	2.45	7.75		
2	O	78.00	7.00	0.50	0.30	2.90	11.30	+ 1.33	1.42
	S	76.98	7.34	0.43	0.29	3.06	11.90		
3	O	73.00	6.00	0.10	0.00	4.90	12.00	+ 0.44	2.61
	S	75.37	6.80	0.09	0.00	5.24	12.50		
4	O	79.00	10.00	0.30	0.00	2.60	8.10	+ 1.73	1.73
	S	79.03	10.24	0.39	0.00	2.59	7.76		
5	O	72.00	11.00	0.40	0.10	6.40	10.10	+ 1.70	1.70
	S	72.11	10.61	0.34	0.08	6.49	10.36		
6	O	68.00	12.00	0.20	0.20	8.90	10.90	+ 1.45	1.45
	S	67.76	12.04	0.00	0.30	9.06	10.85		

Contd....

TABLE 4.7 contd....

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
7	O	61.00	17.00	1.30	0.00	8.30	12.40	+ 1.58	1.58
	S	61.29	16.59	1.38	0.00	8.29	12.44		
8	O	61.00	20.00	1.60	0.50	5.70	11.20	- 6.68	9.14
	S	44.78	28.36	2.24	0.75	8.21	15.67		
9	O	48.00	23.00	1.50	0.00	9.50	18.00	+ 1.00	1.00
	S	47.62	23.81	1.90	0.00	9.52	17.14		
10	O	76.93	12.87	0.00	0.00	5.10	5.10	+ 0.56	0.56
	S	76.92	12.82	0.00	0.00	5.13	5.13		
11	O	52.09	22.91	2.08	1.10	10.40	10.42	-44.36	44.36
	S	33.33	37.50	0.00	4.17	20.83	4.17		
12	O	46.16	15.38	7.69	0.00	23.07	7.70	0.00	0.00
	S	46.15	15.38	7.69	0.00	23.08	7.69		

* Observed trips

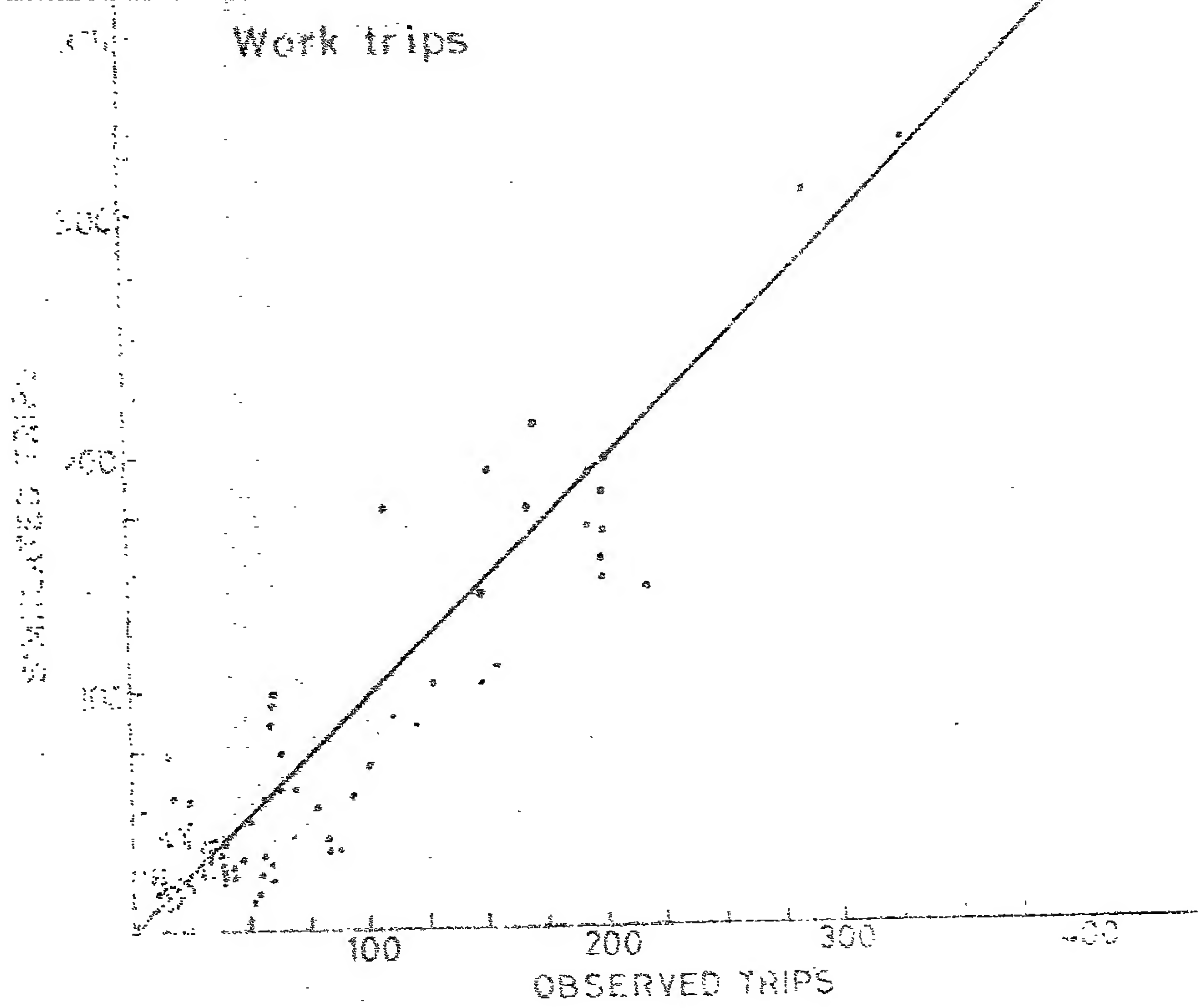
** Simulated trips

4.8.5 Distribution Submodel

The logic of the trip end choice for an individual is interconnected to the status group and mode choice submodels. The decision process is further based upon the employment/education opportunities in different zones,travel time and travel cost. The destination of each simulated individual is identified as per the decision process and the total number of trips by purpose are determined for each O-D pair. These correspond to each element of the O-D matrix. It is highly ambitious to expect from a model that the simulated values of each O-D pair match with the observed ones. Most of the trip distribution models use travel time frequency distribution as the property for empirical validation (Ben-Akiva, 1974; Lerman, 1975).

Figs. 4.23 and 4.24 show the comparison of the observed and simulated trips between major O-D pairs. The points are spread out on both sides indicating the reasonableness of the model. As mentioned earlier this comparison is at a microlevel and cannot by itself be used for any definite conclusions. Some of the aggregated measures that have been used for validation in this study are distribution of employment/education opportunities, travel times and travel costs.

The destination zones are clubbed into 10 groups in the model, each group representing a certain range of attractions i.e., employment and educational opportunities. A comparison of the simulated and observed frequencies of employment /



COMPARISON OF OBSERVED AND SIMULATED WORK TRIPS BETWEEN DIFFERENT O-D PAIRS

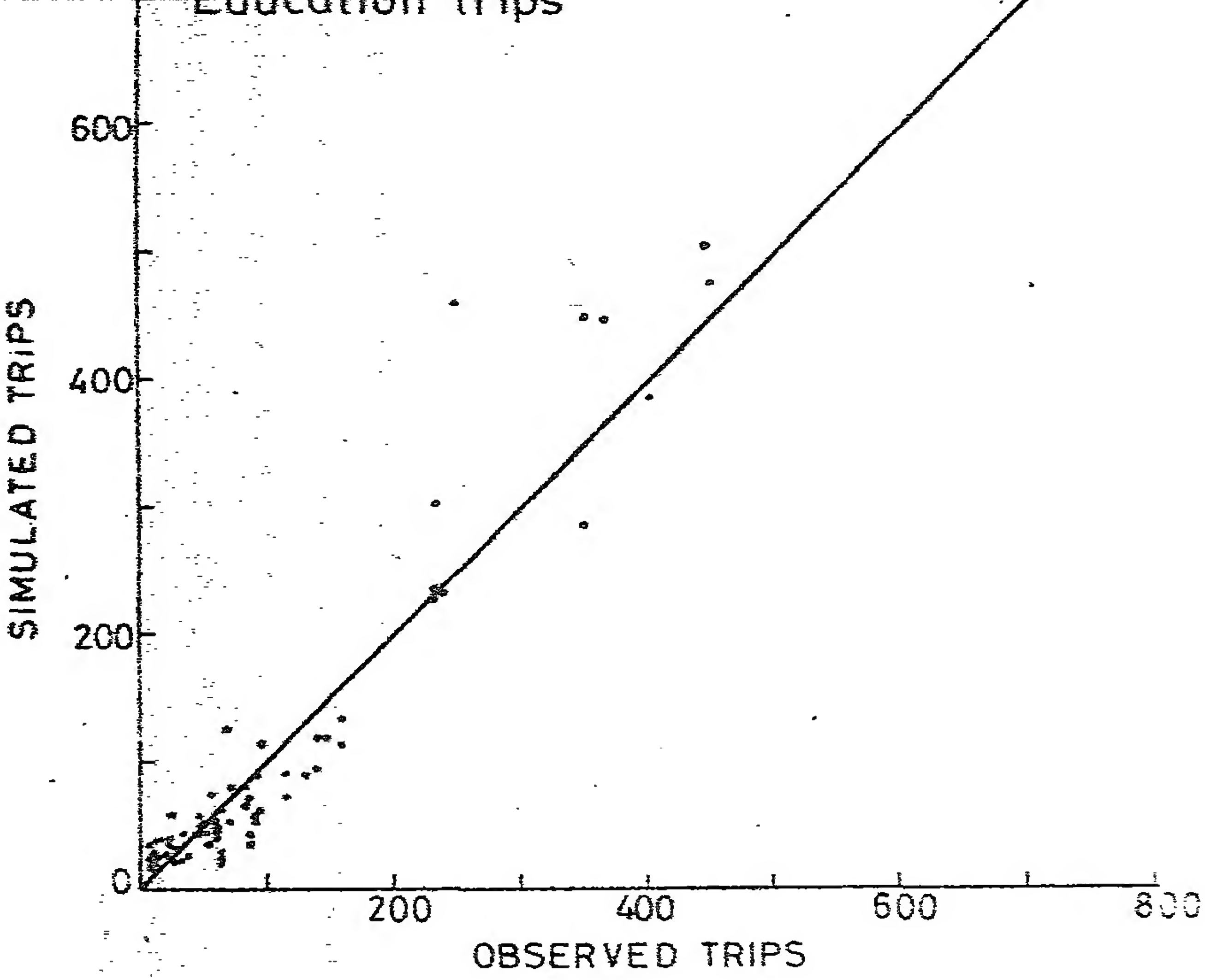


FIG. 4.24 COMPARISON OF OBSERVED AND SIMULATED EDUCATION TRIPS BETWEEN DIFFERENT O-D PAIRS

education opportunities for each of the destination groups is shown in Fig. 4.25. This indicates that the variation is within $\pm$ 10 percent for almost all the cases, the mean percentage difference being -1.48 for work trips and -4.41 for education trips. The cumulative relative frequencies of the observed and simulated trips are shown in Figs. 4.26 and 4.27 for work and education trips respectively. The closeness of the simulated and observed values indicates that the choice of destination group is quite reasonable. As the trip distribution model is interconnected with the status group submodel, a comparison of the observed and simulated frequencies in each of the destination groups is also made for each of the status group and they are shown in Figs. 4.28 and 4.29 for one particular status group. The simulated frequencies compare well with the observed ones thereby further reinforcing the decision logic of the model.

The comparison of the observed and simulated frequencies of trips for each of the travel time groups is shown in Fig. 4.30. Each point in the figure represents one particular travel time group and as mentioned earlier in sub-section 4.4.3 that there are 10 travel time groups of different ranges. The comparison indicates that the values compare well within $\pm$ 10 percent. The close nature of the observed and simulated values are also indicated in the cumulative relative frequencies shown in Figs. 4.31 and 4.32. The observed and simulated frequencies of different travel time groups are also compared for each of the status groups and they are shown in Figs. 4.33

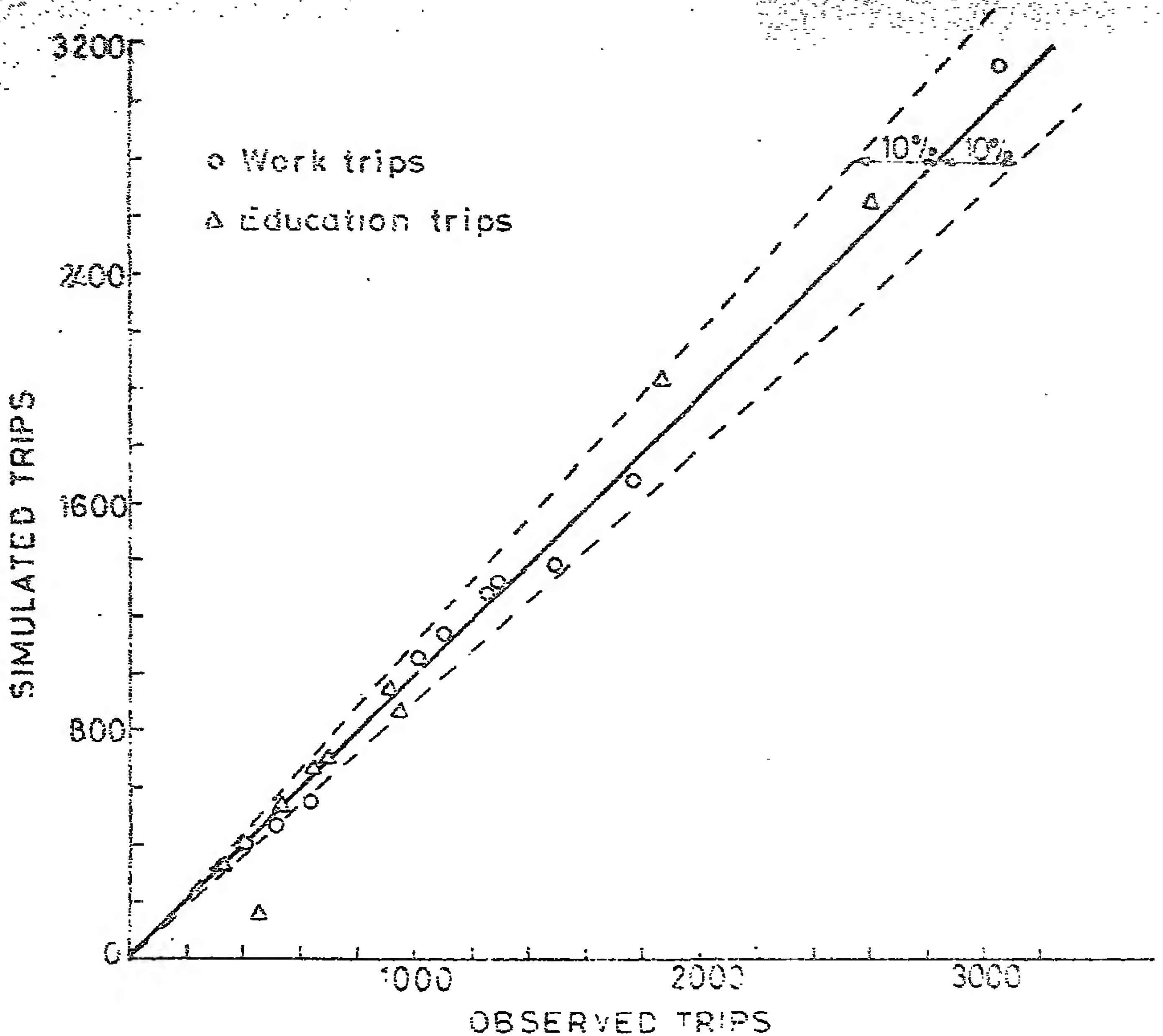


FIG.4.25 COMPARISON OF OBSERVED AND SIMULATED TRIP ATTRACTIONS TO VARIOUS DESTINATION GROUPS

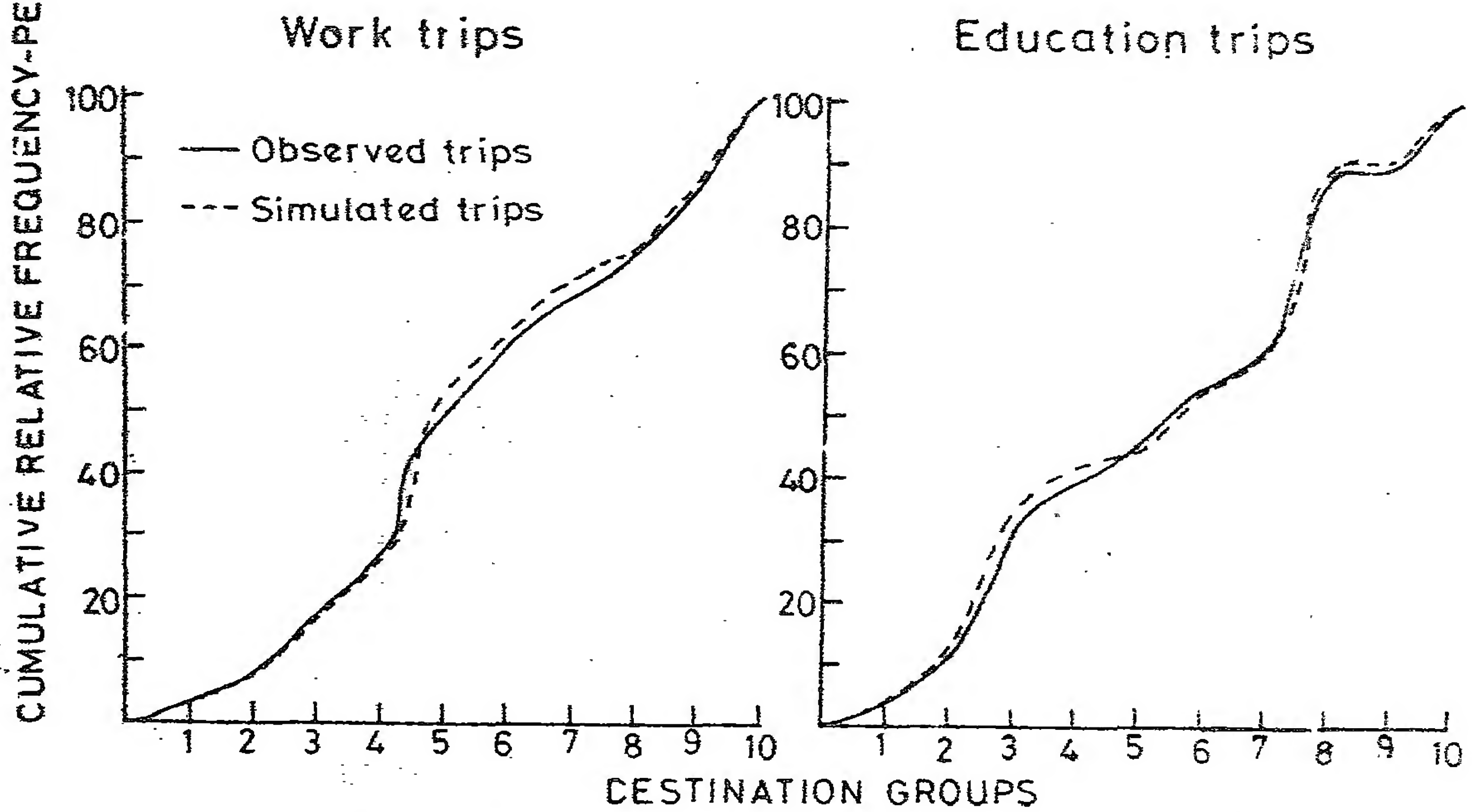


FIG. 4·26 FIG. 4·27

CUMULATIVE FREQUENCY DISTRIBUTION OF TRIPS BY DIFFERENT DESTINATION GROUPS

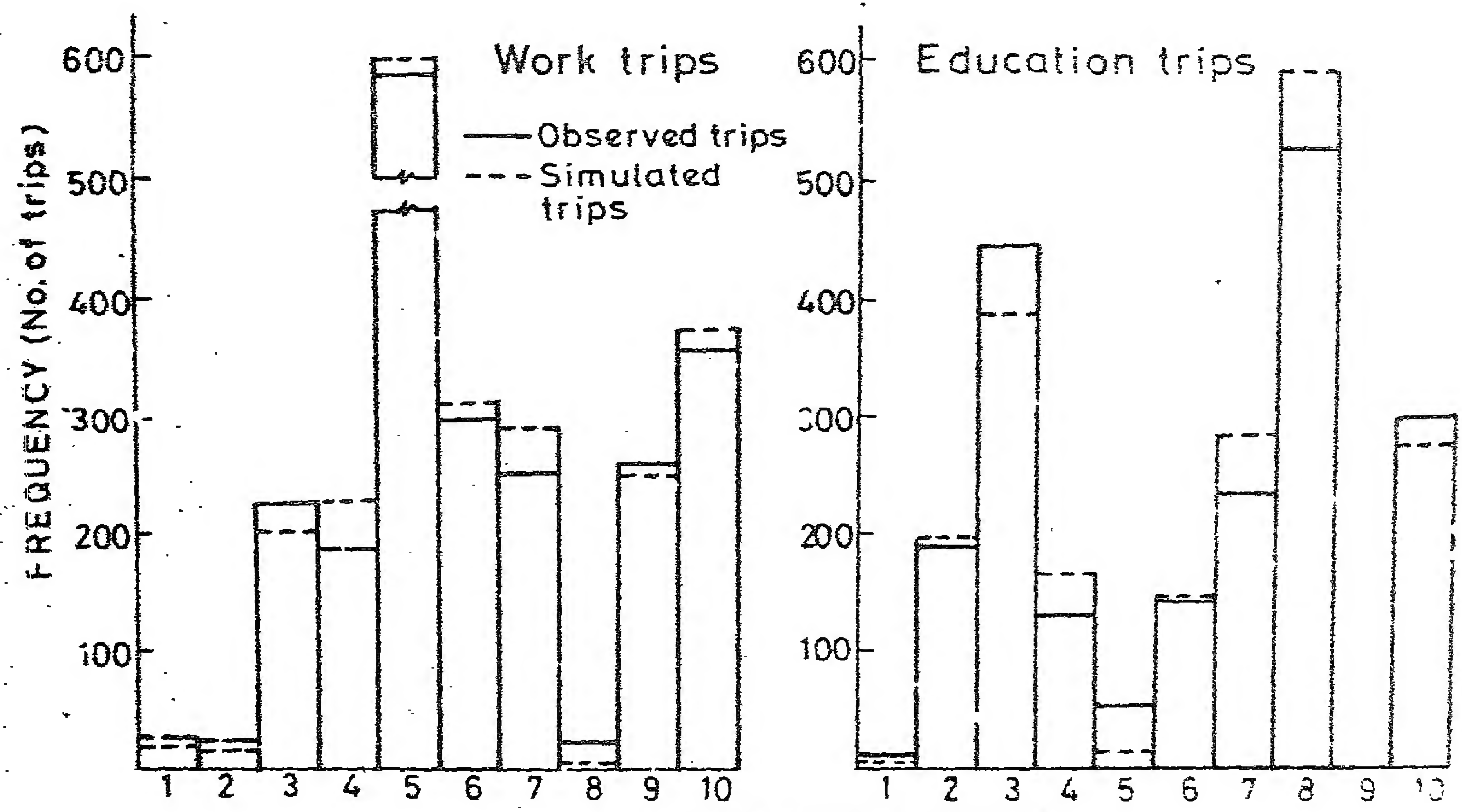


FIG. 4.28 FIG. 4.29

FREQUENCY DISTRIBUTION OF TRIPS BY DIFFERENT DESTINATION GROUPS FOR STATUS GROUP 1

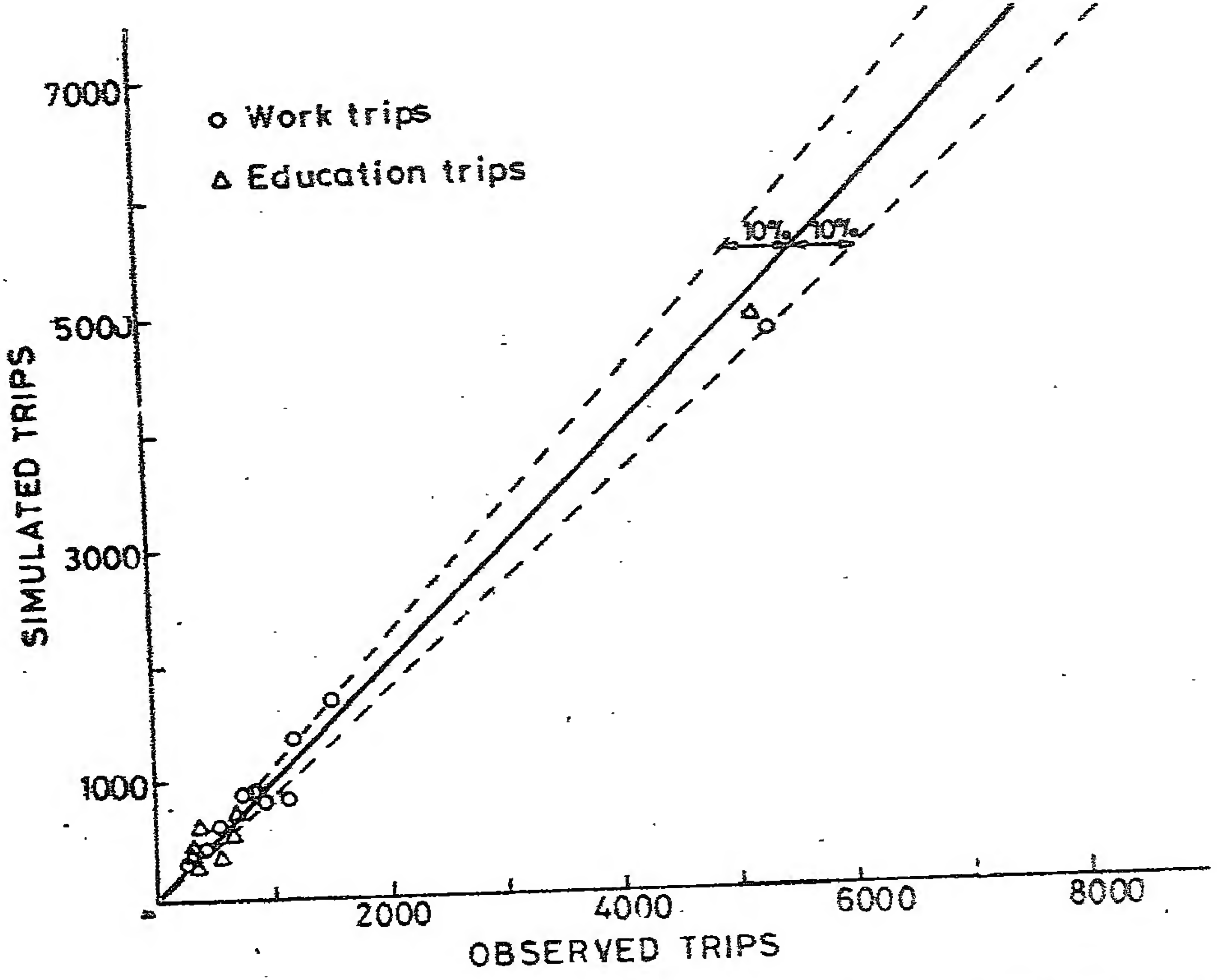


FIG-4-30 COMPARISON OF OBSERVED AND SIMULATED TRIPS BY DIFFERENT TRAVEL TIME GROUPS

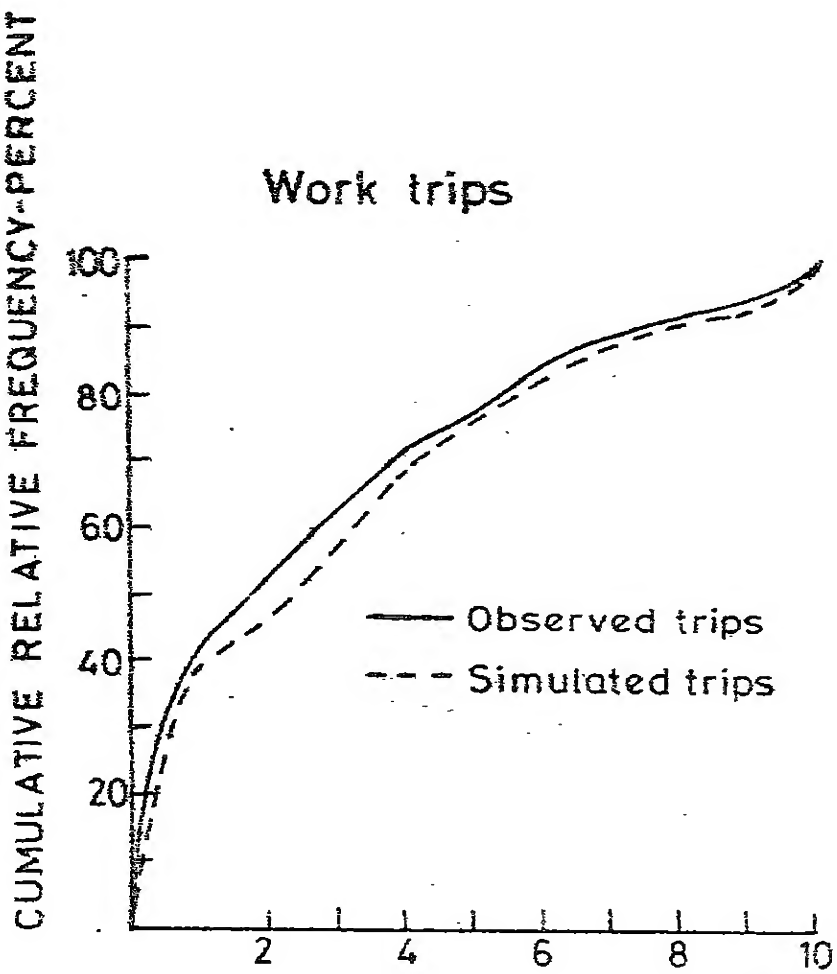


FIG. 4·31

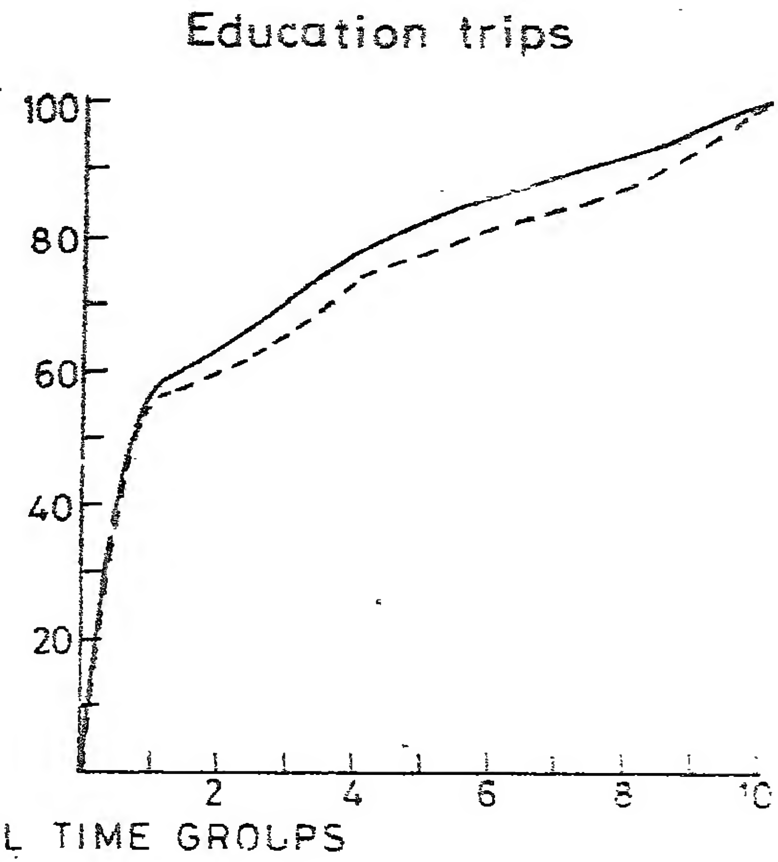


FIG. 4·32

CUMULATIVE FREQUENCY DISTRIBUTION OF TRIPS BY DIFFERENT TRAVEL TIME GROUPS

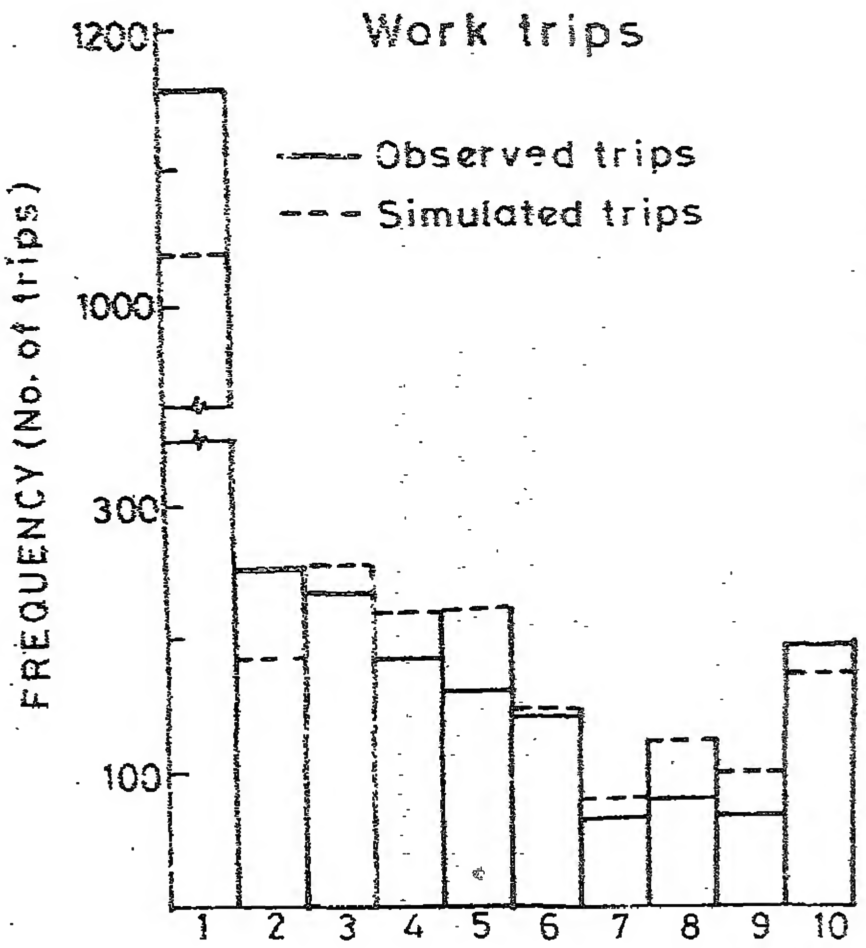


FIG. 4.33

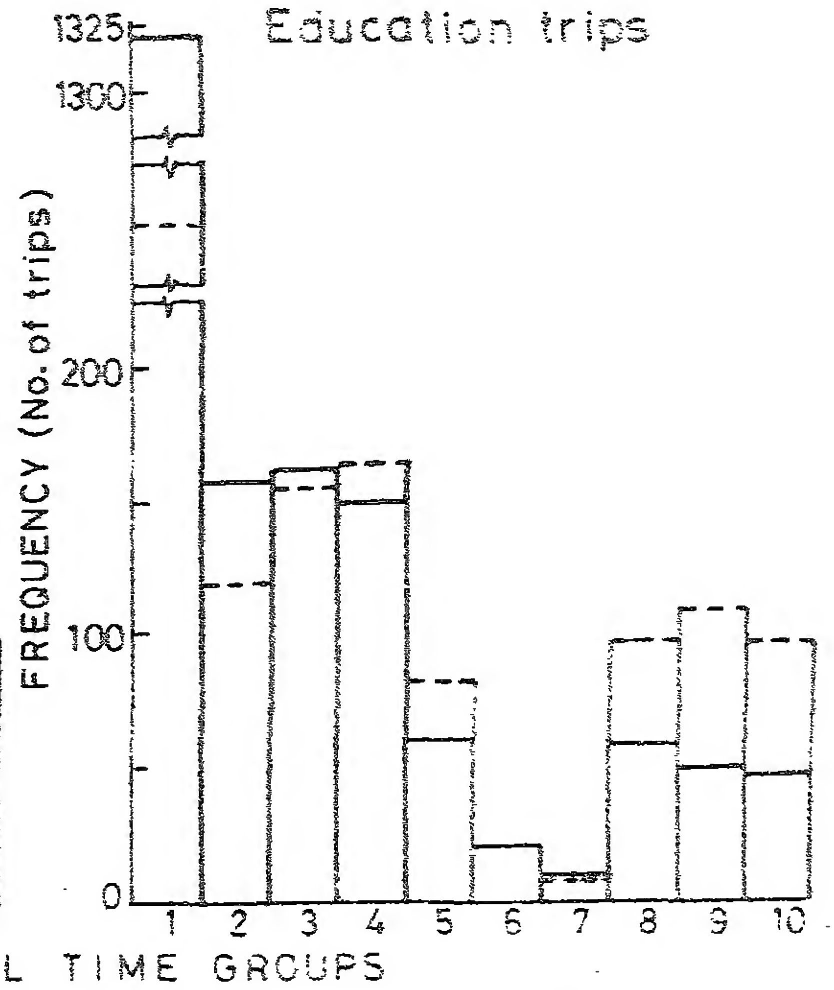


FIG. 4.34

FREQUENCY DISTRIBUTION OF TRIPS BY DIFFERENT TRAVEL TIME GROUPS FOR STATUS GROUP 1

and 4.34 for one particular status group. The above comparisons indicate that the simulated values are quite close to the observed ones with regard to travel time, thereby indicating that the simulation model is a sufficiently good representation of the system. The average trip length by each status grous is also calculated and a comparison of the observed and simulated values is shown in Figs. 4.35 and 4.36. It is observed that, in general, travel time difference is within ± 2 minutes for most of the status groups. Large variations within 3 to 6 minutes is observed generally for individuals residing between intermediate and outer cordons (status groups 3,6,9) in both cases. This is quite understandable as the sample size is low in these cases.

Fig. 4.37 shows the comparison of the observed and simulated frequencies of trips for each of the seven travel cost groups. The comparison indicates that the values compare well within ± 10 percent when the size of the sample is large. But for small sample size, the differences are high. The nature of observed and simulated values are also indicated in the cumulative relative frequencies shown in Figs. 4.38 and 4.39 for work and education trips. The observed and simulated frequencies of different travel cost groups are also compared seperately for each of the status groups and are shown in Figs. 4.40 and 4.41 for one particular case. The comparisons indicate that the simulated values are closed to the observed ones with regard to travel cost, indicating that the simulation model is a sufficiently good representation of the system. The average

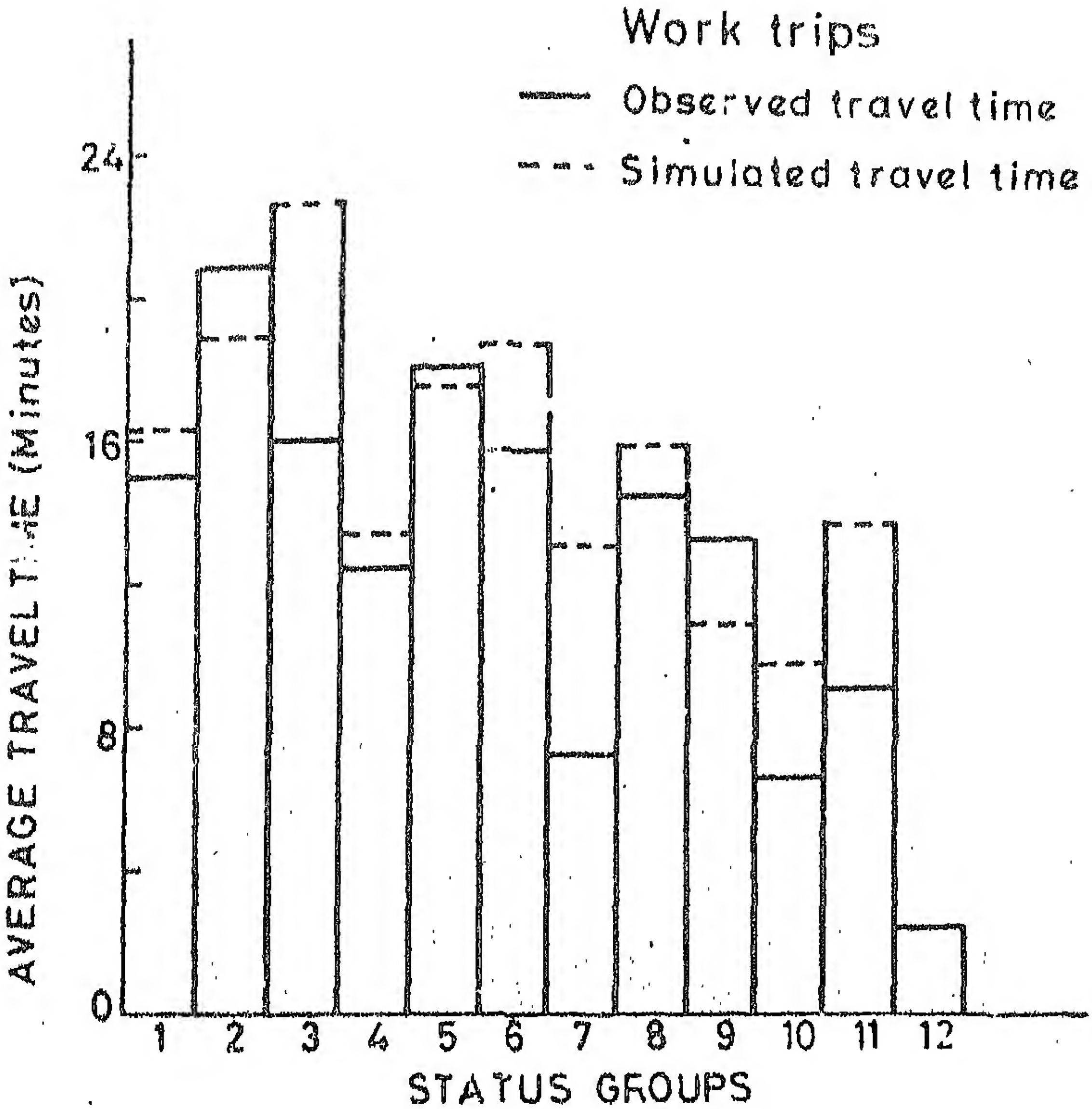


FIG.4.35 FREQUENCY DISTRIBUTION OF TRAVEL TIMES FOR DIFFERENT STATUS GROUPS

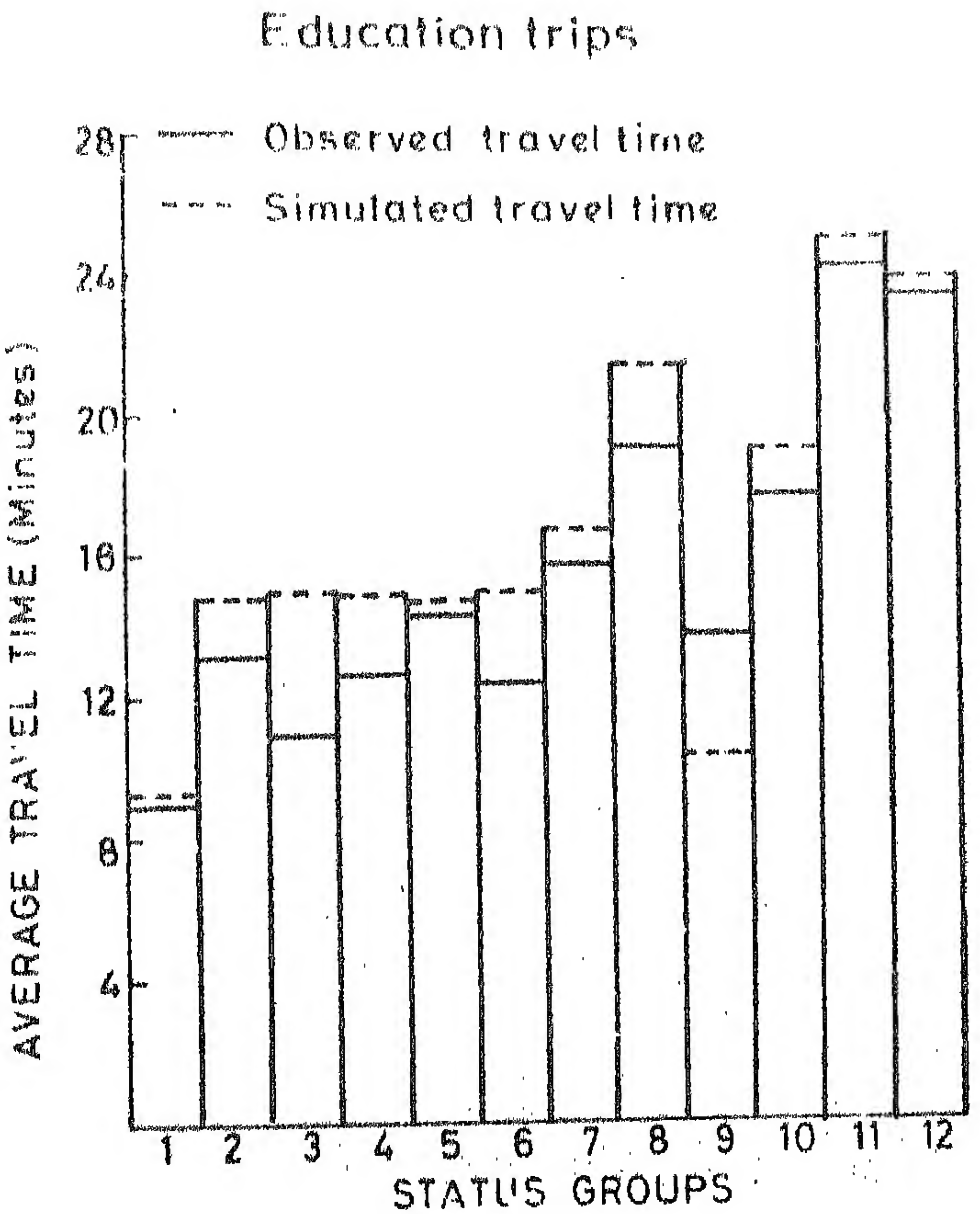


FIG. 4.36 FREQUENCY DISTRIBUTION OF TRAVEL TIMES FOR DIFFERENT STATUS GROUPS

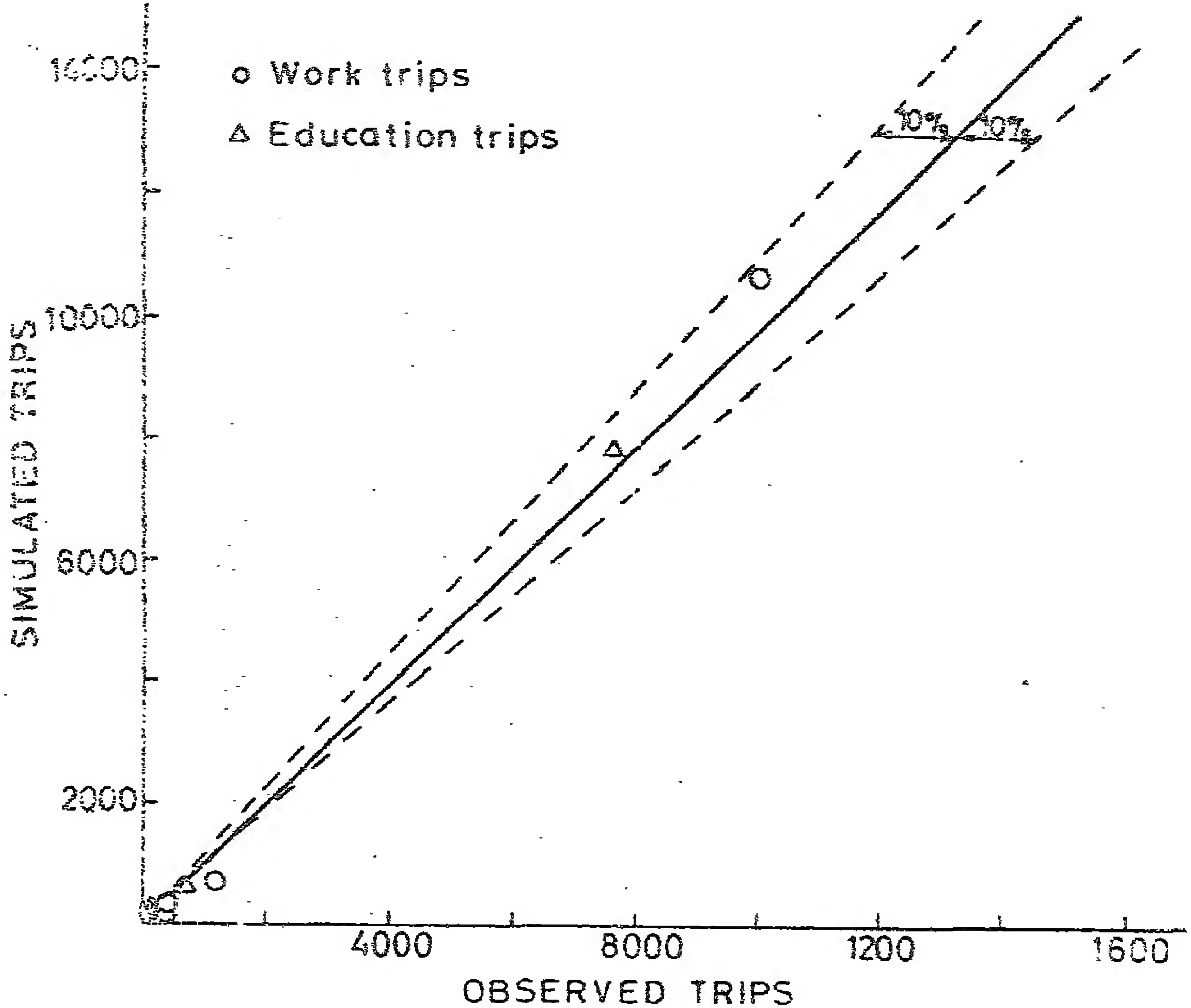


FIG. 4.37 COMPARISON OF OBSERVED AND SIMULATED TRIPS BY DIFFERENT TRAVEL COST GROUPS

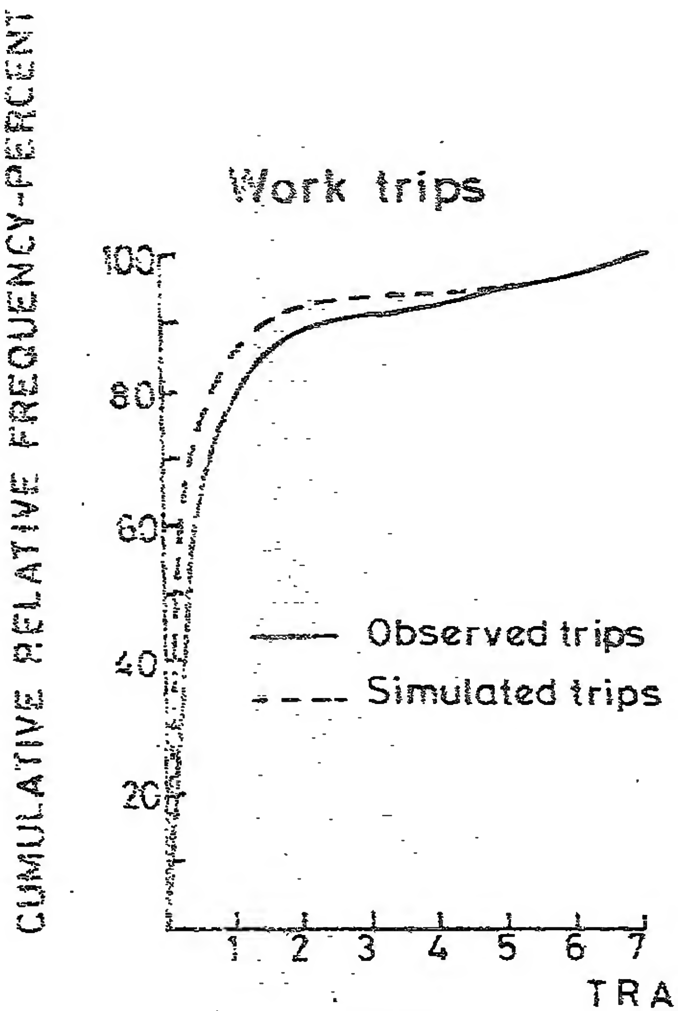


FIG-4·38

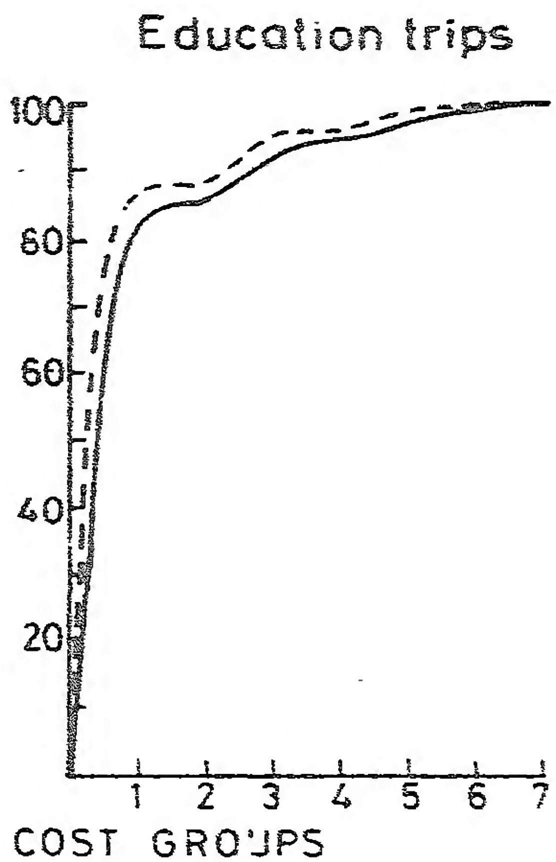


FIG-4·39

CUMULATIVE FREQUENCY DISTRIBUTION OF TRIPS BY DIFFERENT TRAVEL COST GROUPS

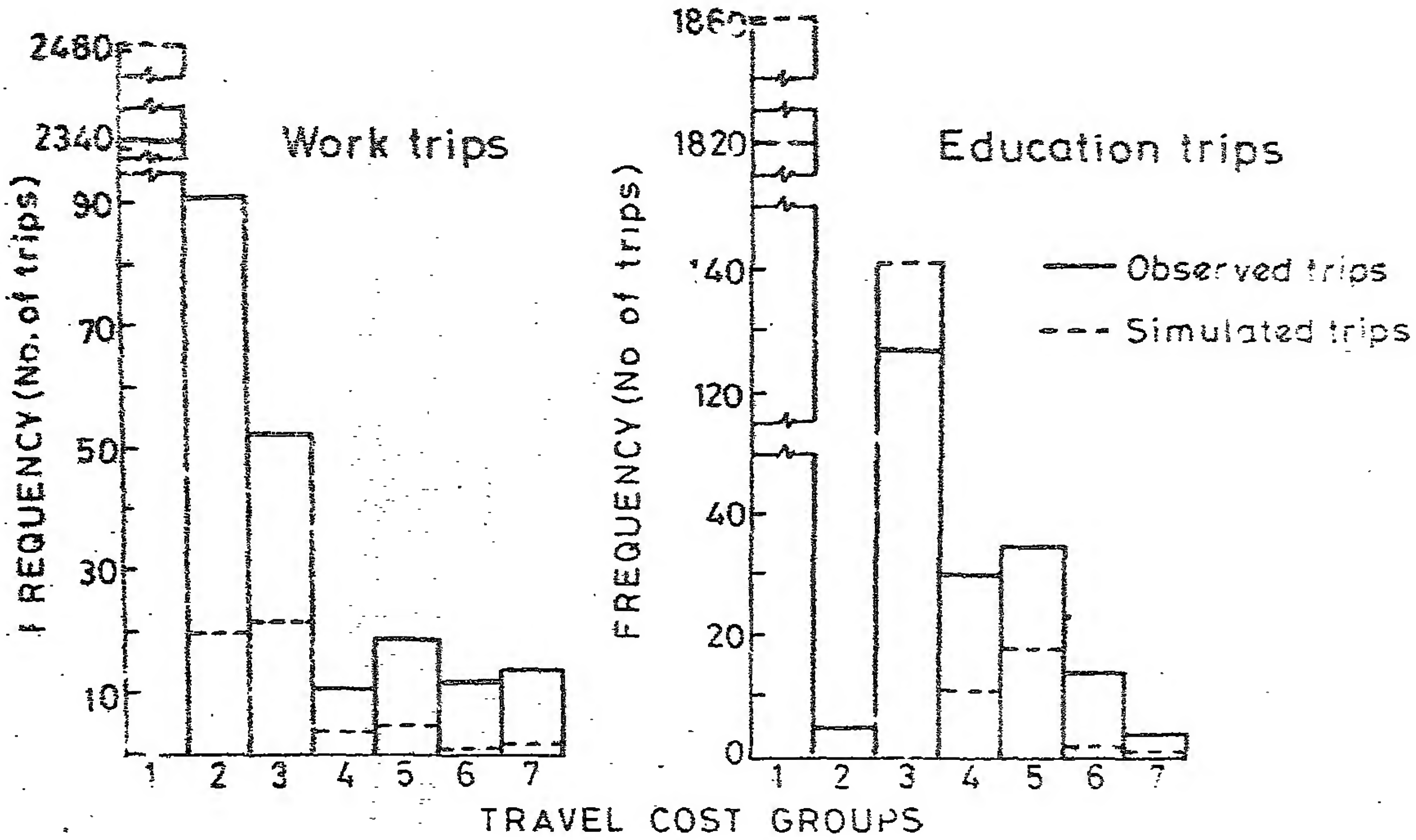
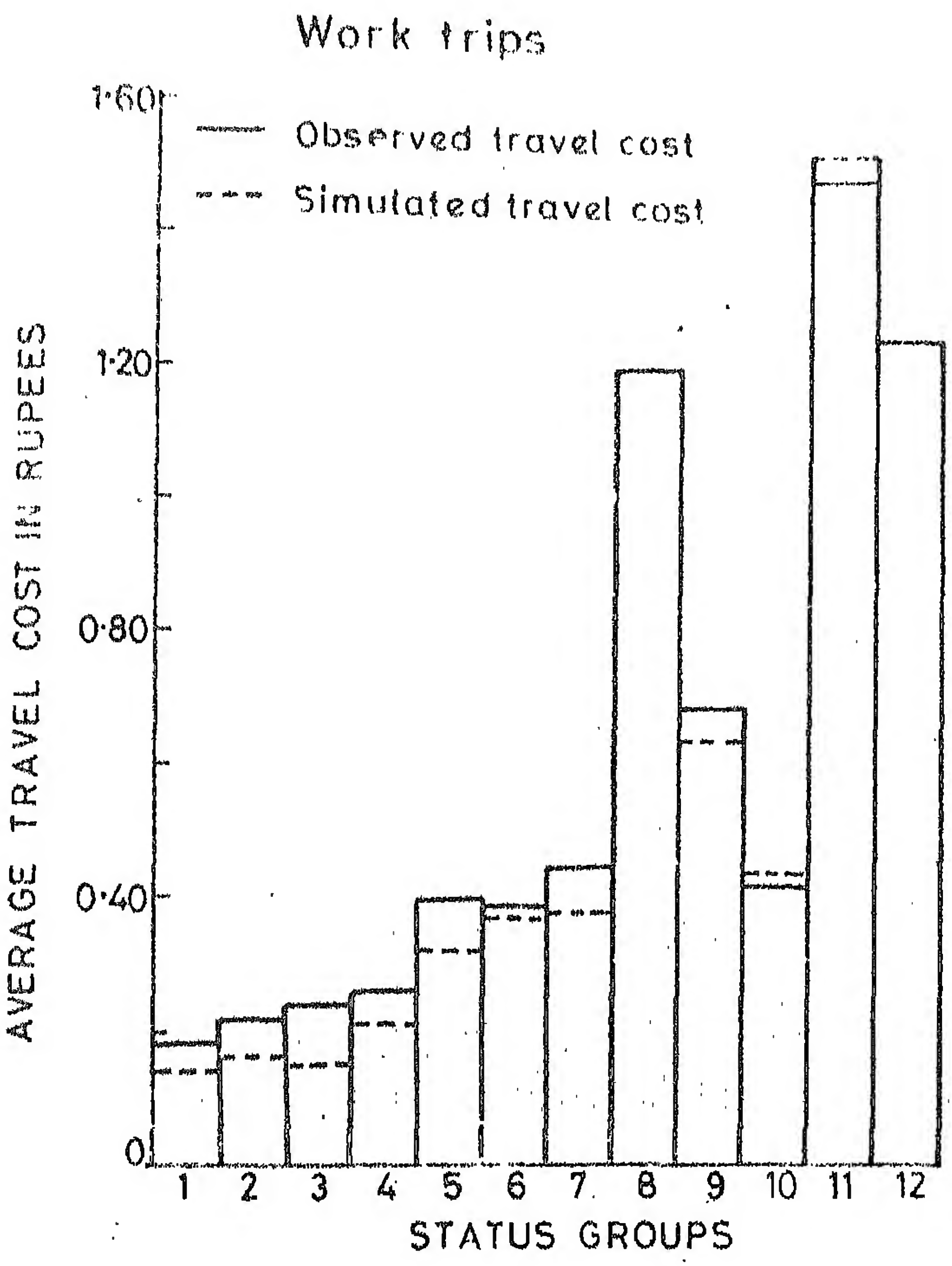


FIG-4-40 FIG-4-41

FREQUENCY DISTRIBUTION OF TRIPS BY DIFFERENT TRAVEL COST GROUPS FOR STATUS GROUP 1

cost of the trips performed for each status group are calculated and a comparison of the observed and simulated values is shown in Figs. 4.42 and 4.43. The results show that the model stands validated with respect to both the modal split and trip distribution aspects of the planning process.

The choice of trip end is based on generated value of travel time and travel cost for the individual. These values are derived from the decision profiles of the individuals status group. Another procedure which is considered on experimental basis is to choose the travel time and travel cost of the individual from the decision profiles of the selected mode. This means that after selecting the mode of travel, the travel time and travel costs are derived only with regard to that mode. This is of more relevance in cities like Kanpur where there are large variations in the travel times and travel costs of the available multiple modes. Based upon this strategy the individuals are simulated and their destinations are selected. The simulated trip distributions obtained by this process also compare well with the observed ones. The relative frequency distributions of the travel times and travel costs for the simulated individuals are shown in Figs. 4.44 to 4.47 along with the observed frequencies and the simulated frequencies as derived from the previous experiment(where the travel time and travel cost are generated from the decision profile of the status group). The results indicate that both the simulation experiments give values close to the observed ones.



FIG·4·42 FREQUENCY DISTRIBUTION OF TRAVEL COSTS FOR DIFFERENT STATUS GROUPS

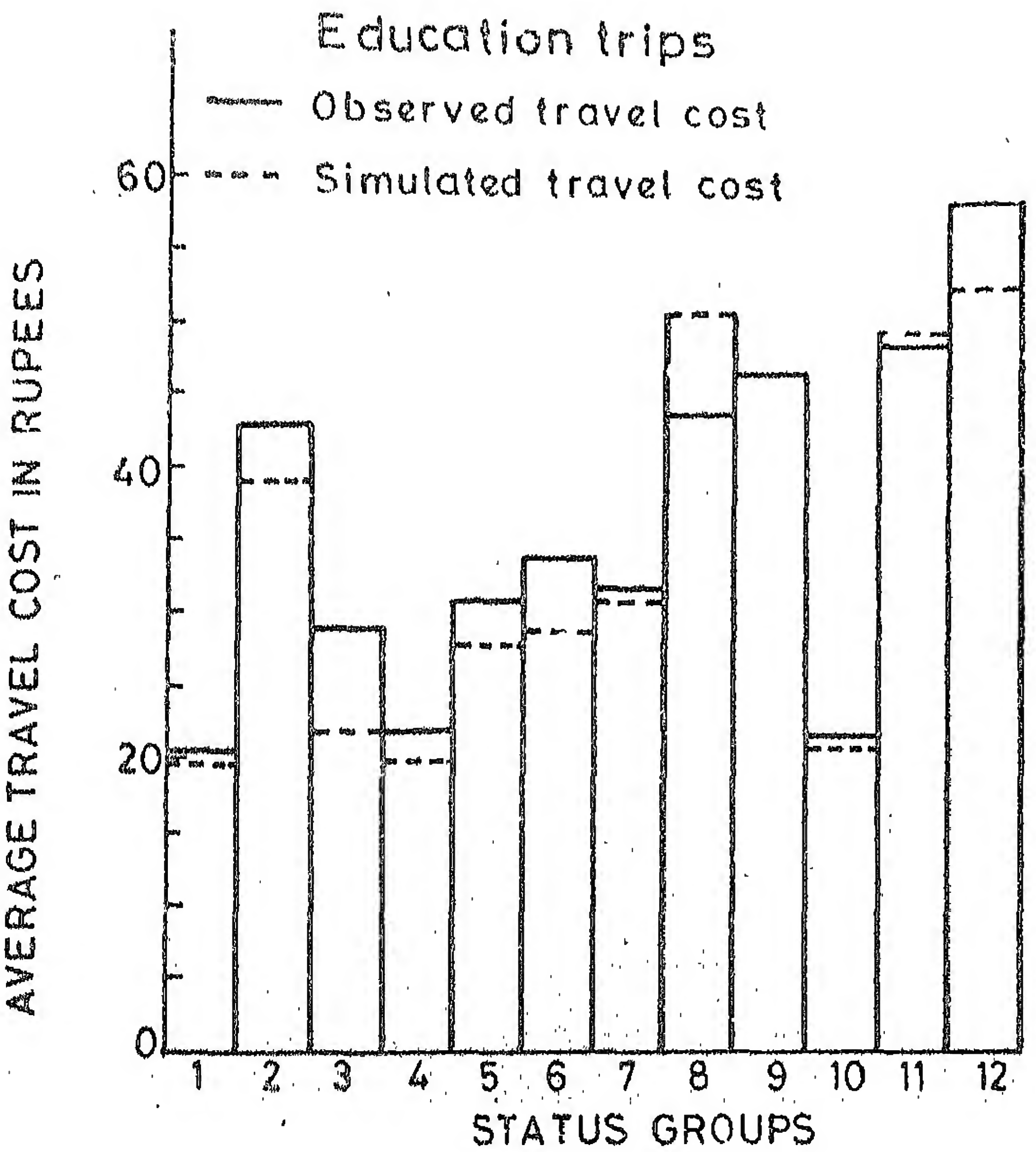


FIG.4·43 FREQUENCY DISTRIBUTION OF TRAVEL COSTS FOR DIFFERENT STATUS GROUPS

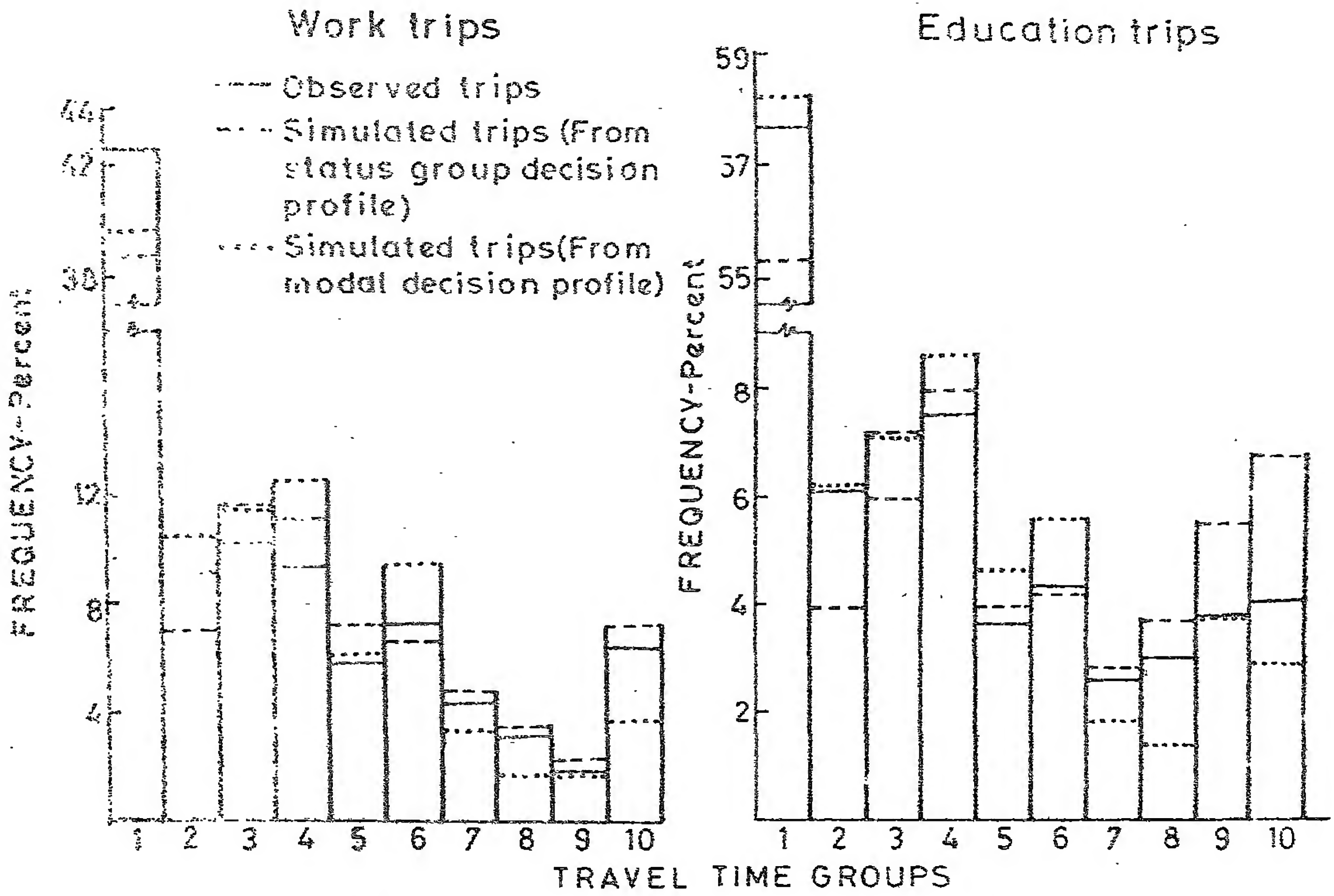


FIG.4.44

FIG.4.45

FREQUENCY DISTRIBUTIONS OF TRAVEL TIME GROUPS

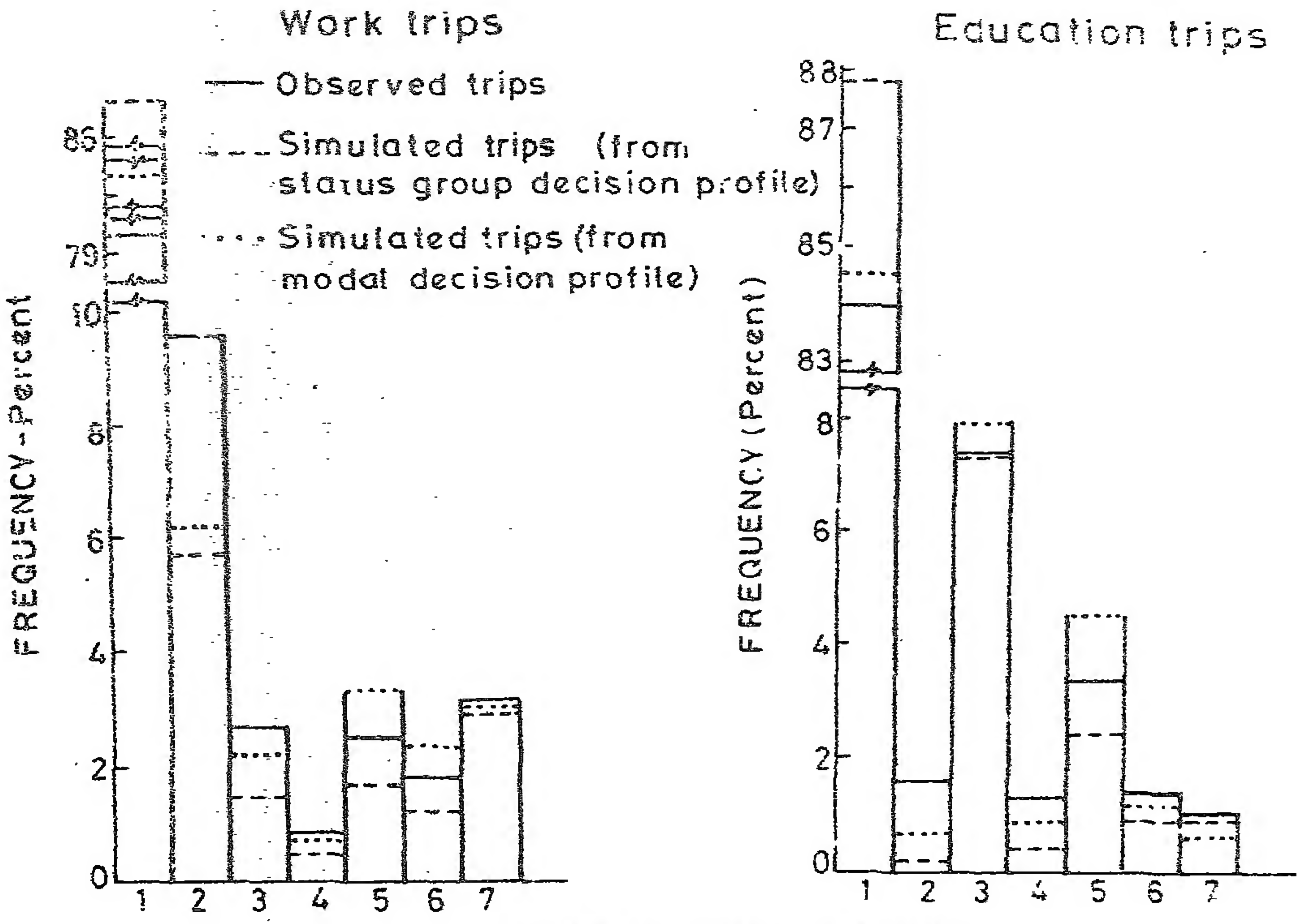


FIG.4.46

FIG.4.47

FREQUENCY DISTRIBUTIONS OF TRAVEL COST GROUPS

It may be desirable that both the alternatives of deriving travel time/travel cost be tested for some other Cities also so as to make some definite conclusions.

4.8.6 Conclusions

The model simulates the modal split and trip distribution as an interconnected individual decision process which is stochastically determined by the location of the individual and the decision profile of his/her status group. The empirical validation of the model indicates its capability of simulating the decision process of the individuals in a complex multimodal situation for a typical large size Indian city. The algorithm is quite sensitive and also flexible enough to incorporate alternative decision processes as is done in this model by considering two strategies of selecting the travel time/travel cost of the individual. The method of defining status groups with homogeneous travel behaviour offers the possibility of testing temporal and spatial stability of travel behaviour pattern. The model is computationally quite efficient as compared to the various types of available programming models. The model may be extended to include the traffic assignment aspect also.

5 PROGRAMMING MODEL FOR COMBINED TRIP DISTRIBUTION AND ASSIGNMENT

5.1 INTRODUCTION

Transport planning process as usually carried out consists of a number of stages. The trip distribution stage is concerned with the estimation of the number of trips per unit time which will be made between each pair of zones in the study area. The estimated pattern of trips depends on the cost of travel between the various pairs of zones and these costs have usually been calculated from fixed costs associated with links of the network. However the link cost increases with the amount of traffic using the link and this is taken into account at the traffic assignment stage when the trip demands obtained from the trip distribution model are allocated to routes through the network. The link costs which correspond to the final estimated traffic flows are, however, not the same as assumed in the trip distribution stage. This problem can be overcome by combining trip distribution and traffic assignment into one stage and describing it by one model. The nature of the generally used seperate models for trip distribution and traffic assignments is briefly discussed in this section along with some of the combined distribution and assignment models. Finally a combined trip distribution and assignment model developed in this study along with its application to the city of Kanpur is presented.

5.2 REVIEW OF LITERATURE

5.2.1 Trip Distribution Model

Several kinds of trip distribution models have been forumulated (Frater, 1954; Brokke, 1958; Martin, Memmott and Bone, 1961; Murchland, 1966; Wilson,1967; Beckmann,1971). The present study is restricted only to the widely used class of gravity models. While the gravity model approach is labelled as being over simplistic, its widespread use is a testimony to the fact that something better is yet to be invented. The number of trips between a given pair of zones in a typical urban area depends on the land use activity in the two zones and the relative ease of interzonal travel afforded by the particular transportation system being considered. In the gravity type distribution model, land use activity is modelled by the estimated number of trips produced in and attracted to each zone. The performance of the transportation network is introduced through a friction factor function. The seperation or friction factor is usually a function of cost of travel between the zones whereby cost is meant by some combination of the travel time, the distance travelled, and the direct monetory cost. Different kinds of gravity models are appropriate to different circumstances and the most important distinctions relate to the availability of estimates for the number of trips which begin and end in each zone, and to the form of cost or deterrence function. Bureau of Public Roads (1965) has developed a technique for calibrating the form of the gravity model for establishing a relationship between travel time factor and travel time for

base year conditions. The solution to the gravity model has also been obtained by various optimization techniques (Murchland, 1966; Wilson, 1967; Tomlin-Tomlin, 1968; Evans, 1970; Beckmann, 1971). Wilson (1970) pioneered the use of entropy models in the study of spatial interaction. Spatial interaction models such as gravity model are based on the entropy maximization principle of information theory. Entropy models are most commonly used to find the most probable numbers of pairings, t_{ij}, between location i and j, given the numbers O_i of homes in location i and D_j of jobs in location j for all locations.

Entropy models are also emerging as valuable tools in the study of various social problems of spatial interaction. Increased flexibility in the models can be obtained by allowing certain constraints to be relaxed from equality to inequality (Jefferson and Scott, 1979). To provide a better understanding, form of the entropy models have also been analysed by geometric programming. Erlander (1977) has suggested an alternative way of interpreting the entropy maximizing approach when used to derive the gravity model formulation for trip distribution. He views the optimization model as a planning model and for this purpose examines two alternative formulations, which are referred to as efficiency and accessibility problems. This interpretation of entropy maximizing framework as a planning model has been questioned by Bruzelius (1981). Hansen (1972), Neuburger (1973), Cochrane (1975) and Nijkamp (1975) have stressed the close links between the entropy approach

and the utility maximization approach as used in conventional microeconomic consumer demand theory. In this they have argued that Wilson's entropy measure should simply be viewed as one specification of a utility function.

5.2.2 Traffic Assignment Model

The purpose of traffic assignment is to provide an estimate of the amount of traffic which will use each part of the transport network under certain conditions. This is generally done by taking the estimated trip demands from the trip distribution model and allocating them to routes through the network according to some hypothesis. The two fundamental traffic assignment principles enunciated by Wardrop (1952) are user-optimizing assignment and system - optimizing assignment. Blunden (1956) and Taylor (1968) found that Wardrop's second criterion describes the route choices of motorists better than Wardrop's first criterion. However, this evidence is far from conclusive and most of the traffic assignment techniques used on a route basis in urban transportation studies assume that Wardrop's first criterion governs the route choices of motorists.

In the recent past, significant progress has been made in the development of efficient algorithms that produce equilibrium flows in assignment of traffic based on Wardrop's first principle. Some of the developed algorithms by Bruynooghe et al. (1968), Dafermos (1971), Leventhal et al. (1973), Murchland (1969), LeBlanc (1973), Nguyen (1974) assign a fixed O-D trip matrix to a road network and result in equilibrium flows. However, very little material is available on the

application of these methods in actual practice. A most efficient and internally coherent algorithm for traffic assignment could be worthless if the resulting flows do not resemble the actual flows. A very important question that needs lot of attention, but generally ignored by researchers, is whether Wardrop's first principle, that traffic will tend to settle in an equilibrium state where no driver can reduce his journey time by choosing a new route, is a sound behavioural assumption (Florian and Nguyen, 1976).

Bruzellus (1979) has made lot of efforts in trying for a precise relationship between microeconomic consumer demand theory and generalized cost which makes some interesting points regarding user behaviour. Traditionally heuristic assignment techniques (capacity-restraint type, incremental or total loading methods) have been employed for finding a user-optimizing flow pattern. Optimization (mathematical programming) techniques have been used for computing a system-optimizing flow pattern. However, the procedural equivalency of the user-optimizing and system-optimizing assugnment principles (Wardrop, 1952; Dafermos and Sparrow, 1969; Potts and Oliver, 1972; Ruiter, 1974; Steenbrink, 1974; Florian, 1976; Wigan, 1977) should make it possible to use either a mathematical optimization technique or a heuristic loading technique to satisfy either of the two traffic assignment principles by choosing the appropriate form of the travel cost-flow function. This characteristic has been examined in theoretical terms in literature but its importance has not been fully realised and

only limited use of it has been made in traffic assignment related work (Akcelik, 1979). A number of techniques have been developed to solve Wardrop's deterministic-user equilibrium problem which are especially relevant for traffic assignment problems in congested urban areas (LeBlanc et al. 1975; Nguyen, 1974). These techniques can be applied to to networks with unlimited link capacities and more recently have been generalized to apply to more realistic case of networks with finite link capacities (Daganzo, 1977 a,b). For lightly congested networks other technique which are labelled 'stochastic network loading' are also available. The traffic assignment methods require that all possible origins and destinations of trips taking place within a study area be represented as if they are taking place to and from a small set of points or centroids. Each centroid is supposed to represent the location of all trip ends within a given zone and this necessarily misrepresents points located at the edges of the zone (Daganzo, 1980). In order to alleviate this problem one could use smaller zones and more centroids, but existing traffic assignment algorithms cannot efficiently handle many centroids.

5.2.3 Combined Trip Distribution and Assignment Model

The usual models of trip distribution and traffic assignment can be equivalently formulated as optimization problems subject to certain constraints. The mathematical programming techniques have generally been used to establish the propertie of these models and to find their solutions.

The mathematical programming background allows to see how seperate models can be combined into a single model in a consistent way. Once the combined model has been formulated, the mathematical problems of establishing existence and uniqueness of solutions, devising algorithms for solving the models, and proving that these algorithms converge to unique solutions can all be attempted by the suitable mathematical programming techniques. A few attempts have been made to combine trip distribution and traffic assignment. Some of them resolved to repeated use of distribution and assignment computer programs to provide feedback between assignment and trip distribution. This is however very expensive to do and the process need not necessarily converge to an overall solution. A few other authors, notably Bruynooghe (1969), Tomlin (1971), Evans (1973), and Florian et.al., (1974) have proposed single models for the combined distribution and assignment problem.

Bruynooghe (1969) has developed a model based on constrained minimization theory. In this the trips can be assigned to the network in some feasible way, so that all link flows are less than the link capacities. Then they can be so assigned in such a way as to satisfy the principle of equal journey costs. Two algorithms are proposed for solving the general form of equivalent minimization problem. Tomlin (1971) and Scheele (1977) considered the problem of combining a doubly constrained gravity model with exponential cost-function, with a traffic assignment model which seeks a

traffic flow pattern minimizing the total cost of all journeys when link travel costs are constant and each link has an associated capacity. This model is formulated directly as a nonlinear problem and is solved by using the Dantzig-Wolfe decomposition principle. Florian et al. (1974) have developed another algorithm similar to Bruynooghe for solving the minimization problem. The method adopts Frank and Wolfe (1956) algorithm for quadratic programming and falls within the class of feasible direction algorithms of nonlinear programming.

Evans (1976) has developed a model based on equivalent optimization problem. He has considered that the demand for trips is described by a gravity model in which the number of trips per unit time which begin and end in each zone are known and cost function either takes on the value of zero or has the form of negative exponentiation of cost for each origin-destination pair. Secondly it is assumed that the cost of travelling along each link of the transportation network is a known strictly increasing function of the traffic flow on that link such that the cost increases indefinitely as the capacity of the link is approached. Thirdly, it is assumed that the trips are assigned to the network in such a way that the resulting traffic pattern satisfies the selfish equilibrium conditions.

Patnaik (1983) has used the Tomlin's model by maximizing the entropy constraint to evaluate the transit system based on system optimization concept, i.e., minimization of the total travel costs of the bus transit system.

A combined trip distribution, modal split and trip assignment model has been suggested by Florian and Nguyen (1977). The main components of the model are entropy type distribution models for each mode that are linked together via common trip production and attraction constraints, an equilibrium type 'user optimized' route choice according to Wardrop's first principle and shortest route choice of transit travellers on the transit networks. It is shown that when several entropy distribution models are linked with route choice models based on Wardrop's 'user optimized' principle, the mode choice is given by a logit model and the destination choice is given by a 'weighted' logit function.

5.3 MODEL DEVELOPMENT

An attempt is made in this study to derive models that combine trip distribution and traffic assignment into a single stage. This model optimizes the system by minimizing the total cost of the system as a whole. The main components of the model are entropy type trip distribution submodel for each mode, and a system optimized assignment submodel. The combined model is solved by using the decomposition principle. All the constraints of the gravity submodel of the distribution problem are linear except the entropy constraint which is convex in the region defined by linear constraints. The gravity submodel is proposed to be calibrated by making use of Davidon-Fletcher Powell (DFP) method of constrained optimization. The gravity submodel through an iterative process evolves the distribution matrix for each mode within the framework of the constraints

of the given trip generation data, the generalized cost, and existing extent of entropy. The assignment submodel minimizes the total travel cost through system optimization for each mode considering the link costs over the network. The dual costs obtained from the assignment submodel replace the generalized costs in gravity submodel as feedback so as to get the revised trip matrices. These two submodels stop interacting when no further change in objective function is possible. Through the successive solutions of these two submodels, the final assigned trips can be obtained.

5.3.1 Trip Distribution Problem

Given the trip generation data, the trip distribution stage is concerned with estimation of the number of trips per unit time which will be made between each pair of zones in the study area. Suppose the number of zones in which the trips begin is m and number of zones in which the trips end is n. Now let the number of trips per unit time begining in origin zone i be p_i ($i = 1,2,\ldots,m$), and the number of trips per unit time ending in destination zone j be a_j($j=1,2,\ldots,n$). Let t_{ij} be the number of trips from origin i to destination j and c_{ij} be the unit generalized cost (or travel time) from origin i to destination j. The total number of trips made in the area is given by $T = \sum_{ij} t_{ij}$. The travel pattern is completely specified by the trip matrix t_{ij} and the following conditions must be satisfied :

$$\sum_j t_{ij} = p_i, \; i = 1,2,\ldots,m \qquad \ldots (5.1)$$

$$\sum_i t_{ij} = a_j, \; j = 1,2,\ldots,n \qquad .. \; (5.2)$$

$$t_{ij} \geq 0 \text{ all } i \text{ and } j .$$

The objective is to minimize the total travel cost, $\sum_{ij} c_{ij} t_{ij}$, or the mean travel cost $\frac{1}{T} \sum_{ij} c_{ij} t_{ij}$. With either of the above two as the objective function and subject to the constraints (Eqns. 5.1 and 5.2), one arrives at the classical transportation problem in linear programming. The solution of this problem results in no trips being made at all between certain origin-destination pairs. However a transportation problem such that it is possible to reach every destination from every origin is needed, that is certain degree of accessibility is required. This can be achieved by including an entropy constraint. The entropy may be looked upon as a measure of trips over the cells of the trip matrix. The higher the value of the entropy, the more even is the distribution. Since there are many pairings which are compatible with the generally available data, it makes sense to have the choice of the most probable set of pairings. This is the 'Principle of Insufficient Reason' of Laplace and the resulting problem is the maximization of entropy (Jefferson and Scott, 1979), given the data available.

For the distribution which is even,

$$\frac{t_{ij}}{T} = \frac{1}{mn} \qquad .. \; (5.3)$$

and $$E_{max} = - \sum_{ij} \frac{t_{ij}}{T} \log \frac{t_{ij}}{T}$$

$$= - mn \frac{1}{mn} \log \frac{1}{mn}$$

i.e., $$E_{max} = \log mn \qquad .. (5.4)$$

where, E_{max} = maximum entropy.

If the distribution is not even, all journeys are concentrated in one cell, then the minimum entropy, i.e., $E_{min} = 0$. Assuming E as the measure of accessibility to be at least equal to or greater than the existing one, E_0, the following relationship is obtained

$$- \sum_{ij} \frac{t_{ij}}{T} \log \frac{t_{ij}}{T} \geq E_0$$

The trip distribution problem is now formulated as

$$\text{Min } Z = \sum_i \sum_j c_{ij} \, t_{ij} \qquad .. (5.5)$$

subject to $$\sum_j t_{ij} = p_i \qquad .. (5.6)$$

$$\sum_i t_{ij} = a_j \qquad .. (5.7)$$

$$- \sum_{ij} \frac{t_{ij}}{T} \log \frac{t_{ij}}{T} \geq E_0 \qquad .. (5.8)$$

$$t_{ij} \geq 0$$

All the above constraints are linear except the entropy constraint which is convex. The function $- \sum_{ij} \frac{t_{ij}}{T} \log \frac{t_{ij}}{T}$ is strictly convex in the region defined by linear constraints (Evans, 1973). The objective function and all the constraints are all differentiable. These conditions guarantee that a solution to the Kuhn-Tucker stationary point problem also is

a solution to the minimization problem (Scheele, 1977).

5.3.1.1 Choice of entropy constraint

The maximum value of the entropy constraint which can be considered is when production and attraction constraints along with the non-negativity constraint are being satisfied. Because of its strict convexity, there exists a feasible solution. The t_{ij} matrix obtained from the model is considered to give the highest accessibility to the system when $t_{ij} = \frac{p_i a_j}{T}$ for the given marginal total of the matrix. In this case, the cost of travel has no influence on the trip matrix. For the trip matrix t_{ij} of the form

$$t_{ij}(\beta) = O_i \; D_j \; \mathrm{EXP}(-\beta \; c_{ij}) \qquad \text{.. (5.9)}$$

where O_i and D_j are some constants. The value of the entropy is obtained as

$$E(t_{ij}) = \sum_{ij} \frac{t_{ij}}{T} \log \frac{t_{ij}}{T}$$

This maximum value is obtained when the value of β becomes zero as per (Eqn.5.9). It is shown by Evans (1973) that the function $E(\beta) = E(t_{ij}(\beta))$ is a strictly increasing function of β when $\beta > 0$. The value of E_0 in the entropy constraint corresponds to β_0 in the solution of the gravity model. This is a one to one correspondence, as the entropy is strictly decreasing in β. β_0 has also a unique corresponding value of the cost c_0. The value of E_0 in the gravity model calibration is

$$E_0 = - E\,(t_{ij}(\beta_0)) \qquad \text{.. (5.10)}$$

to find the appropriate step length α_i, are employed to find the minimum of the modified objective function. The optimization problem can be stated as finding the vector $\bar{X}$ of the calibration coefficient which minimizes the objective function $F(\bar{X})$, subject to the constraints $g_j(\bar{X}) \leq 0$ $(j=1,2,..,N)$. The methodology of most of the numerical methods of optimization is to produce a sequence of improved approximations to the optimum in the following manner :

$$\bar{X}_{i+1} = \bar{X}_i + \alpha_i^* \bar{S}_i \qquad .. (5.12)$$

where, $\bar{X}_{i+1}$ = vector of design variables at the new point,

$\bar{X}_i$ = vector of design variables at the previous point,

$\bar{S}_i$ = a suitable direction along which the value of the objective function decreases, and

α_i^* = an appropriate step length for movement along $\bar{S}_i$.

Davidon-Fletcher-Powell (DFP) method also known as the variable matrix method has been found to be stable and applicable for a large class of problems. As such this method has been chosen to find out the search direction $\bar{S}$. Since this method makes use of gradient of the objective function, it has been coupled with the cubic interpolation method to find the appropriate step length α. Constraints have been taken care of through exterior penalty function approach. The constrained optimization problem is converted into an unconstrained optimization problem by constructing a function ϕ, using the objective function $F(\bar{X})$, the constraints $g_j(\bar{X})$ and a penalty parameter r as given below

$$\phi_k = \phi(\bar{X}, r_k) = F(\bar{X}) + r_k \sum_{j=1}^{N} g_j(\bar{X})^2 \qquad .. (5.13)$$

where the exponent z is a non-negative constant. The choice of z = 2 is selected (Fox, 1971). If the unconstrained minimization of function ϕ is repeated for a sequence of values of the penalty parameter, r_k (k = 1,2,....), the solution may be brought to converge to that of the original problem. Hence penalty function methods are also known as sequential unconstrained minimization techniques (SUMT). The framework for the method is as follows:

(i) starting with an initial value of $\bar{X}$ with a moderate value of r_k find a vector $\bar{X}_s$ that minimizes the equation 5.13,

(ii) investigate the constraints to determine whether the point $\bar{X}_s$ is in the feasible domain,

(iii) if the result of step (ii) is true, terminate, otherwise pick an $r_k' > r_k$ and starting from $\bar{X}_s$, return to step (i), minimizing $\phi(\bar{X},\bar{r}_k')$.

Once the algorithm has come in a feasible domain, all the subsequent solutions generated will also be within the feasible domain, since the constraint boundries act as barriers during the minimization process.

The constrained nonlinear programming problems of the current study have been converted into unconstrained nonlinear programming problems through the use of exterior penalty functions. Davidon-Fletcher-Powell method (DFP) to find the search direction $\bar{S}$ and cubic interpolation method to find the appropriate step length α, are then employed to find the minimum of the objective function. The steps involved

in the solution of the constrained optimization problem may be summarized as follows:

(i) The solution is commenced from an initial calibration coefficients of $\bar{X}_1$, a suitable value is chosen for the penalty parameter r and counter k set equal to 1.

(ii) The modified objective function $\phi(\bar{X}, r_k)$ is minimized to get $\bar{X}_k^*$ using DFP algorithm and cubic interpolation method as explained in steps (iii) through (viii).

(iii) The algorithm starts with the initial point $\bar{X}_1$ and a M x N positive definite symmetrix matrix $[H]_1$, where M is the number of design variables. The interation number is set equal to 1.

(iv) The gradient of the modified objective function, $\nabla\phi_i$, at the point $\bar{X}_i$ is computed, from which, the search direction $\bar{S}_i$ is found as

$$\bar{S}_i = -[H]_i \nabla\phi_i \qquad .. (5.14)$$

(v) The minimizing step length α_i^* in the direction $\bar{S}_i$ is found, using the cubic interpolation method in four stages. First, the search direction $\bar{S}_i$ is normalized so that a step size $\alpha = 1$ is acceptable. Then, the directional derivatives of the function ϕ are used to establish bounds on α^* as the slope has to change from a negative value to a positive value, if the minimum has been bracketed. In the third stage, an approximate value of α^* is found by representing $\phi(\alpha)$, in the bounded interval, by a cubic polynomial. This cubic polynomial is refitted in the fourth stage, if

the value of α^* found in the third stage does not satisfy the convergence criteria. A better point in the design space is then found as

$$\bar{X}_{i+1} = \bar{X}_i + \alpha_i^* \bar{S}_i \quad .. (5.15)$$

(vi) The new point $\bar{X}_{i+1}$ is tested for optimality. If $\bar{X}_{i+1}$ is optimal, the iterative procedure is terminated.

(vii) Otherwise, the $[H]_i$ matrix is updated as

$$[H]_{i+1} = [H]_i + [M]_i + [N]_i \quad .. (5.16)$$

where, $M_i = \dfrac{\bar{S}_i \, \bar{S}_i^T}{\bar{S}_i^T \, \bar{Q}_i}$,

$$N_i = \frac{([H]_i \, \bar{Q}_i)\,([H]_i \, \bar{Q}_i)^T}{\bar{Q}_i^T \, [H]_i \, \bar{Q}_i}, \quad .. (5.17)$$

and $$\bar{Q}_i = \nabla\phi_{i+1} - \nabla\phi_i \quad .. (5.18)$$

(viii) The new iteration number i is set equal to i+1 and the new iteration commenced from step (iv).

(ix) X_k^* is tested for optimality. If it is optimal the process is terminated,

otherwise the penality parameter is modified as $r_{k+1} = c r_k$, where c is greater than 1.

(x) The new value of k is set equal to k+1, the new starting point $\bar{X}_1$ set equal to $\bar{X}_k^*$ and the next minimization cycle is commenced from step (ii).

Calibration of parameters A and β

The revised value of friction factor f_i' is obtained as

$$f_i' = f_i \frac{OD_i}{GM_i} \quad .. (5.19)$$

where, f_i = the value of friction factor for the ith interval as used in calibration just completed,

f_i' = the value of friction factor for the ith interval to be used in the next iteration,

OD_i = the percentage of total trips occuring for ith travel time/cost interval observed from the data,

GM_i = the percentage of total trips occuring for ith travel time/cost interval obtained from the model,

and NT = number of travel time/cost intervals.

The objective is to minimize the sum of the squares of the difference between the values of friction factors used in the calibration just completed and the value to be used in the next iteration , i.e. ,

$$\text{Minimize} \quad F = \sum_{i=1}^{NT} [f_i - f_i']^2 \qquad .. (5.20)$$

The constraint is that the difference between the observed trip lengths and the trip lengths obtained from the model should be equal to $\pm$ 3 percent of the observed trip lengths. Therefore, the constraint can be written as ,

$$\frac{\left[\dfrac{ABS\ (TOB_i - TGM_i)}{TOB_i}\right]}{0.03} - 1 \leq 0 \qquad .. (5.21)$$

where, TOB_i = the observed trips for each travel time interval,

and TGM_i = the trips observed from the model for each travel time interval.

This turns out to be a mathematical problem of minimization which can be handled easily by optimization technique, dealing

with nonlinear constrained optimization problem. When the above constraint is satisfied the value of A and β are taken as the final calibration coefficients.

5.3.1.3 Solution of the trip distribution problem

The distribution problem is formulated as

$$\text{Minimize} \quad Z = \sum_i \sum_j c_{ij} \, t_{ij} \qquad .. (5.22)$$

subject to

$$\sum_i t_{ij} = p_i \qquad .. (5.23)$$

$$\sum_j t_{ij} = a_j \qquad .. (5.24)$$

$$-\sum_{ij} \frac{t_{ij}}{T} \log \frac{t_{ij}}{T} \geq E_0 \qquad .. (5.25)$$

$$t_{ij} \geq 0 \qquad .. (5.26)$$

Constraints (Eqns. 5.23, 5.24 and 5.26) are linear while entropy constraint (Eqn. 5.25) is convex. Further the function is strictly convex in the region defined by linear constraints. The objective function and the constraints are also differentiabie (Evans, 1973). These conditions guarantee that a solution to the Kuhn-Tucker stationary point problem also is a solution to the above minimization problem. The cost and entropy both are decreasing functions of β, as such in the optimal solution

$$-E(t_{ij}(\beta)) = E_0 \qquad .. (5.27)$$

Since $-E(t_{ij}(\beta))$ is strictly decreasing, the above equation in the scalar variable β has a unique solution. To start with, β is replaced by β_0 which has been estimated from the trip matrix. The function values $-E(t_{ij}(\beta))$ are calculated as follows :

As the first approximation of $t_{ij} = T_{ij}$, the following relationship is assumed

$$T_{ij} = A \operatorname{Exp} (-\beta \, c_{ij}) \qquad .. (5.28)$$

Using the Furness or cross factor techniques as described by Martin et al. (1961), the prescribed marginal totals, p_i and a_j, are obtained by successive adjustments to row sums, T_{i*}, and column sums, T_{*j}. The iterative process is briefly described in the following steps:

(i) by making a summation over each row, $T_{i*} = \sum_j T_{ij}$ is obtained,

(ii) new T_{ij} values are computed as

$$T_{ij}^{new} = \frac{T_{ij}^{old} \, p_i}{T_{i*}} ;$$

(iii) summation over each column gives $T_{*j} = \sum_i T_{ij}$ and

(iv) new T_{ij} values are computed through

$$T_{ij}^{new} = \frac{T_{ij}^{old} \, a_j}{T_{*j}} ;$$

(v) if marginal totals are nearly obtained or within the specified tolerance, then the iterative procedure stops ;else it is repeated i.e., go to step (i); the tolerance used is approximately the one which is given by Marksjo (1970)

$$p_{max} = \max_i \frac{p_i}{T_{i*}} ; \qquad p_{min} = \min_i \frac{p_i}{T_{i*}}$$

$$a_{max} = \max_j \frac{a_j}{T_{*j}} ; \qquad a_{min} = \min_j \frac{a_j}{T_{*j}}$$

The algorithm stops when,

$$p_{max}\ a_{max} \quad \text{and} \quad \frac{1}{p_{min}\ a_{min}} < 1 + \frac{1}{6\,T}$$

The tolerance can also be taken as (1 ± 0.03) instead. This means that there exists an $\varepsilon > 0$, such that,

$$\left| \sum_j T_{ij}(\beta) - p_i \right| < \varepsilon \quad i = 1,2,..,m$$

$$\left| \sum_i T_{ij}(\beta) - a_j \right| < \varepsilon \quad j = 1,2,..,n$$

Thus after having obtained the solution $T_{ij}(\beta)$, the corresponding entropy $- E(T_{ij}(\beta))$ is obtained.

5.3.2 Combined Distribution and Assignment Problem

5.3.2.1 Problem formulation

Traffic assignment phase provides an estimate of the flow t_{ij} to be assigned to various links in the network. The two fundamental traffic assignment principles enunciated by Wardrop are 'user-optimizing assignment' and 'system optimizing assignment'. Some of the combined distribution and assignment problems (Bruynooghe, 1969; Potts and Oliver, 1972; Evans, 1973; Dafermos, 1974; and Florian, Nguyen and Ferland, 1975) are based on user optimizing principle. But an important question is whether the user optimizing principle, that traffic will tend to settle in an equilibrium state when no driver can reduce his journey time by choosing a new route, is a sound behavioural assumption. This model primarly based on Indian city conditions, uses the system optimizing formulation of the problem. The road network is represented mathematically by a directed graph consisting of collection of vertices and links. The vertices represent zones centroids or major

intersections. A two way road is represented by two oppositely directed links. To start with it is assumed that the flow from i to j, t_{ij}, is given:

Let, I_p = link p, p = 1,2,....,NL

y_p = capacity of the link p,

c_p = cost per unit flow on link p,

(i,j,q)= path number q between i and j,

NL = total number of links,

TN_{ij} = total number of paths between i and j,

$$a^{ij}_{pq} = \text{the link incidence matrix},\quad = \begin{cases} 1 & \text{if } I_p \in (i,j,q) \\ 0 & \text{otherwise} \end{cases}$$

The cost per unit flow on any path (i,j,p) is thus,

$$\sum_{p=1}^{NL} c_p \; a^{ij}_{pq}$$

Let x^{ij}_q be the flow along the path q from i to j. Then the assignment problem for the given t_{ij} can be formulated for minimizing the total cost as

$$\text{Minimize} \quad Tc = \sum_i \sum_j \sum_q \sum_p c_p \; a^{ij}_{pq} \; x^{ij}_q \qquad .. \ (5.29)$$

$$\text{subject to} \quad \sum_i \sum_j \sum_q a^{ij}_{pq} \; x^{ij}_q \leq y_p \,, \quad p = 1,2,\ldots, NL \qquad .. \ (5.30)$$

$$\sum_q x^{ij}_q = t_{ij} \text{ for all } i,j \qquad .. \ (5.31)$$

$$x^{ij}_q \geq 0 \qquad q = 1,2,\ldots,TN_{ij}, \ i=1,2,\ldots,m; \ j=1,2,\ldots,n$$

The above assignment problem for the given t_{ij} can easily be combined with the distribution problem of sub-section 5.3.1

by adding the entropy, the production and the attractions constraints and allowing t_{ij} to vary.

The resulting formulation of the combined distribution and assignment problem with the variables as t_{ij} and x_q^{ij} can be written as

$$\text{Minimize} \quad Tc = \sum_i \sum_j \sum_q \sum_p c_p \, a_{pq}^{ij} \, x_q^{ij} \qquad .. (5.32)$$

$$\text{subject to} \quad \sum_i \sum_j \sum_q a_{pq}^{ij} \, x_q^{ij} \leq y_p \,, \quad p = 1,2,..,NL \qquad .. (5.33)$$

$$\sum_q x_q^{ij} = t_{ij}, \text{ all } i,j \qquad .. (5.34)$$

$$\sum_j t_{ij} = p_i, \quad i = 1,2,....,m \qquad .. (5.35)$$

$$\sum_i t_{ij} = a_j \,, \quad j = 1,2,...,n \qquad .. (5.36)$$

$$-\sum_{ij} \frac{t_{ij}}{T} \log \frac{t_{ij}}{T} \geq E_0, \text{ for all } i,j \qquad .. (5.37)$$

This problem is essentially a network flow problem with capacity constraints given by Eqn.(5.33), flow requirements given by Eqn.(5.34), along with the marginal totals of the origin-destination matrix as also the accessibility requirements of the flow matrix (Eqn. 5.37).

The above formulated problem is linear except for one constraint which is the accessibility requirement. This constraint is strictly convex. The feasibility region for t_{ij} is thus convex and is bounded because of production and attraction constraints and $t_{ij} > 0$. An equivalent formulation

of the problem is

$$\text{Minimize} \quad TC = \sum_i \sum_j \sum_q \sum_p c_p \, a^{ij}_{pq} \, x^{ij}_q \qquad .. (5.38)$$

$$\text{subject to} \quad \sum_i \sum_j \sum_q a^{ij}_{pq} \, x^{ij}_q \leq y_p, \quad p = 1,2,..,NL \qquad .. (5.39)$$

$$\sum_q x^{ij}_q - \lambda t_{ij} = 0 \text{, all a,b} \qquad .. (5.40)$$

$$\lambda = 1 \qquad .. (5.41)$$

$$x^{ij}_q \geq 0 \text{, all i,j,q,} \quad \lambda \geq 0 \qquad .. (5.42)$$

$$\sum_j t_{ij} = p_i \text{, all i} \qquad .. (5.43)$$

$$\sum_i t_{ij} = a_j \text{, all j} \qquad .. (5.44)$$

$$- \sum_{ij} \frac{t_{ij}}{T} \log \frac{t_{ij}}{T} \geq E_0 \qquad .. (5.45)$$

$$t_{ij} \geq 0 \text{, all i, j} \qquad .. (5.46)$$

The relations 5.38 to 5.42 can be looked upon as a linear program in the variables x^{ij}_q and λ, with variable coefficients t_{ij}. t_{ij} belong to the convex, bounded and closed region defined by the constraints 5.43 to 5.46. This combined problem of distribution and assignment may be reformulated into two equivalent formulations, namely, assignment submodel and gravity submodel, the successive solutions of which solve the whole problem.

Assignment submodel

$$\text{Minimize} \quad TC_A = \sum_i \sum_j \sum_q \sum_p c_p \, a^{ij}_{pq} \, x^{ij}_q \qquad .. (5.47)$$

subject to $\sum_i \sum_j \sum_p a_{pq}^{ij} \; x_q^{ij} \leq y_p$ .. (5.48)

$$\sum_q x_q^{ij} - \sum_k \lambda_k \; t_{ij}^k = 0 , \quad \text{all } a,b \qquad .. (5.49)$$

$$\sum_k \lambda_k = 1 \qquad .. (5.50)$$

$$x_q^{ij} \geq 0 , \; \lambda_k \geq 0 \qquad .. (5.51)$$

where, λ_k is the variable in addition to x_q^{ij} ,λ_k has t_{ij} as its variable coefficients.

When the Lagrange function is formed,

$$L_g = \sum_i \sum_j \sum_q \sum_p c_p \; a_{pq}^{ij} \, x_q^{ij} + \sum_p \pi_p (\sum_i \sum_j \sum_q a_{pq}^{ij} \, x_q^{ij} - y_p)$$

$$+ \sum_i \sum_j \alpha_{ij} (\sum_k \lambda_k \; f_{ij}^k - \sum_q x_q^{ij}) + \delta (1 - \sum_k \lambda_k) \qquad .. (5.52)$$

The constraints of the dual problem are obtained by differentiating Eqn. (5.52) as,

$$\frac{\partial L_g}{\partial x_q^{ij}} = \sum_p c_p \; a_{pq}^{ij} + \sum_p \pi_i \; a_{pq}^{ij} - \alpha_{ij} \geq 0 \qquad .. (5.53)$$

$$\frac{\partial L_g}{\partial \; k} = \sum_i \sum_j \alpha_{ij} \; t_{ij}^k - \delta \;\; \geq 0 \qquad .. (5.54)$$

It is assumed now that $(\bar{x}_q^{ij}, \bar{\lambda}_k)$ and $(\bar{\pi}_p, \bar{\alpha}_{ij}, \bar{\delta})$ are optimal solutions to the assignment submodel and its dual respectively. Then a new point t_{ij}^k giving an optimum value of the objective function can be found if there is a trip matrix t_{ij} such that ,

$$\sum_i \sum_j \bar{\alpha}_{ij} \; t_{ij} - \bar{\delta} \; \leq \; 0 \qquad .. (5.55)$$

Gravity submodel

$$\text{Minimize } TC_G = \sum_i \sum_j \bar{\alpha}_{ij} t_{ij} \quad .. \quad (5.56)$$

subject to

$$\sum_i t_{ij} = p_i \, , \text{ all } i \quad .. \quad (5.57)$$

$$\sum_i t_{ij} = a_j \, , \text{ all } j \quad .. \quad (5.58)$$

$$-\sum_{ij} \frac{t_{ij}}{T} \log \frac{t_{ij}}{T} \geq E_0 \quad .. \quad (5.59)$$

$$t_{ij} \geq 0$$

5.3.2.2 Solution of the combined distribution and assignment problem

The combined distribution and assignment problem has been reformulated into equivalent formulations of two submodels, namely, assignment submodel and gravity submodel. A successive solution of these submodels produces a solution of the whole problem. The iterations are described by the following algorithm.

Step 0 (i) Initial values α_{ij}^0, δ^0 are chosen. The value of α_{ij}^0 can be taken as per the shortest path algorithm, and δ^0 can be set to ∞.

Step s (ii) For $\alpha_{ij}^s = \alpha_{ij}^{s-1}$, the gravity submodel is solved. Let the solution be t_{ij}^s.

(iii) If $\sum_i \sum_j \alpha_{ij}^{s-1} t_{ij}^s < \delta^{s-1}$, a column is added to the new point t_{ij}^s in the assignment submodel and go to step (iv); otherwise stop.

(iv) Solve the assignment submodel which is a linear program of the form,

$$\text{Minimize} \quad TC_A = \sum_i \sum_j \sum_q \sum_p c_p \, a_{pq}^{ij} \, x_q^{ij} \qquad .. \ (5.60)$$

$$\text{subject to} \quad \sum_i \sum_j \sum_q a_{pq}^{ij} \, x_q^{ij} \le y_p \quad p=1,2,..,NL \qquad .. \ (5.61)$$

$$\sum_i \sum_j \sum_q a_{pq}^{ij} \, x_{pq}^{ij} \, x_q^{ij} \le y_p \quad p = 1,2,..,NL \qquad .. \ (5.61)$$

$$\sum_q x_q^{ij} - \sum_{l=1}^{k} \lambda_l \, t_{ij}^{l} = 0 \ , \ \text{for all } i,j \qquad .. \ (5.62)$$

$$\sum_{l=1}^{k} \lambda_l = 1 \qquad .. \ (5.63)$$

$$x_q^{ij} \ge 0 \ , \ \lambda_l \ge 0$$

Its dual can be formulated as

$$\text{Maximize} \quad TC_{AD} = - \sum_p \pi_p \, y_p + \delta \qquad .. \ (5.64)$$

$$\text{subject to} \quad \sum_p (c_p + \pi_p) \, a_{pq}^{ij} - \alpha_{ij} \ge 0, \ \text{for all } p,i,j \qquad .. \ (5.65)$$

$$\sum_{ij} \alpha_{ij} \, t_{ij}^{l} - \delta \ge 0 \quad l = 1,2,...,s \qquad .. \ (5.66)$$

$$\pi_p \ge 0 \ , \ p = 1,2...,NL$$

The dual values π_{ij}^{s} and δ^{s} are to enter (ii) and (iii) to proceed to (iv) if needed.

The linear program to solve the assignment submodel is of considerably large size, both in terms of number of variables and number of constraints. Solution of the linear program is just one step in the iterative procedure. To reduce the size of the problem only significant O-D pairs may be considered. To further reduce the number of variables only two or at the most three alternative paths between each pair may be considered. The shortest path is always considered along

with one or two alternative paths that are next to the shortest one.

When a solution to the linear programming problem is obtained, then the possibility of allowing more alternative paths to enter can be checked by calculating the reduced cost for all variables. The reduced cost for the variable x_q^{ij} is

$$\sum_p c_p \, a_{pq}^{ij} + \sum_p \pi_p \, a_{pq}^{ij} - \alpha_{ij} \qquad .. \ (5.67)$$

Let the dual solution obtained be $(\bar{\pi}_p, \bar{\alpha}_{ij}, \bar{\delta})$ for the set of paths considered. If $\sum_p (c_p + \bar{\pi}_p) \, a_{pq}^{ij} < \bar{\alpha}_{ij}$ for some path, then this path should enter the set of alternative paths. The paths of least reduced cost is one with the smallest value of $\sum_p (c_p + \bar{\pi}_p) \, a_{pq}^{ij}$. This is the shortest path between i and j. When the link costs are greater than $\bar{\alpha}_{ij}$, then a solution to problem iv has been obtained, as it is not possible for any of the remaining paths to reduce the objective function.

5.4 MODEL APPLICATION

The combined distribution and assignment model as formulated in sub-section 5.3.2.1 has been tested on the city of Kanpur as a pilot implementation of the algorithms developed. This pilot application brought clearly into focus the various issues associated with the model as presented in the following sections.

5.4.1 Strategy for Model Application

Typical Indian cities like Kanpur have multimodes of travel available on the roads. Further the data of various

socio economic and travel aspects as available are limited. The strategy for model application has to be well planned.

Given the traffic generation, the model performs the trip distribution and traffic assignment phases by minimizing the system travel cost. Kanpur city has in general six major modes of travel . These modes have significant variations in operational characteristics. A comprehensive study to estimate the travel cost by different modes on the roads has still not been carried out. In the absence of the related information and various other constraints, travel time is taken as a measure of travel cost in this model. The Decision Process Models of Chapter 2 perform the modal split after the trip generation. It is planned that, after the modal split, the distribution and assignment phase be carried out seperately for each mode of travel. This will result in system optimization within each mode.

The model is applied both for work and education trips for which the household survey data are available. The Kanpur road network as shown in Fig. 2.1 indicates that in making a trip between an O-D pair not many alternative paths are available. This is because of very limited number of links in the network. In some of the situations only one path may be available. The choice of path depends upon the travel time/travel cost as the level of service is associated with the links of the path. Motor vehicles while choosing a particular path do consider the road conditions, traffic interactions, obstructions like intersections etc., on the alternative paths. Slower vehicles like bicycles

and cycle rikshaws are generally not affected by the traffic interactions of other vehicles, though they themselves significantly affect the flow characteristics of motor vehicles. These slow vehicles consider distance as a major measure in choosing a path and thus generally always opt for the shortest path. Bicycle and cycle rikshaw trips are distributed as per the gravity submodel, but due to the above mentioned constraints it is proposed to do away with the assignment submodel and these trips are assigned to the shortest paths. But for the motor vehicular modes (car,bus, motorcycle/scooter) distribution and assignment aspects are carried out as per the model both for work and education trips.

The model requires an input of the alternative available paths between each O-D pair, and estimates the flow on each of these paths. In this study only two alternative paths for each O-D pair are considered. These include the shortest travel time path and the one with next higher travel time. For some of the O-D pairs only one path had to be specified when no other alternative is available. The network as considered in this study (Fig. 5.1) contains 312 unidirected links, and the alternative paths for 236 significant O-D pairs are listed in Appendix II.

5.4.1.1 Link capacities

The model includes the link capacity constraint while assigning the trips. The capacities of each of the links need to be specified. As in this study the distribution and assignment phase are being done seperately for each mode the link

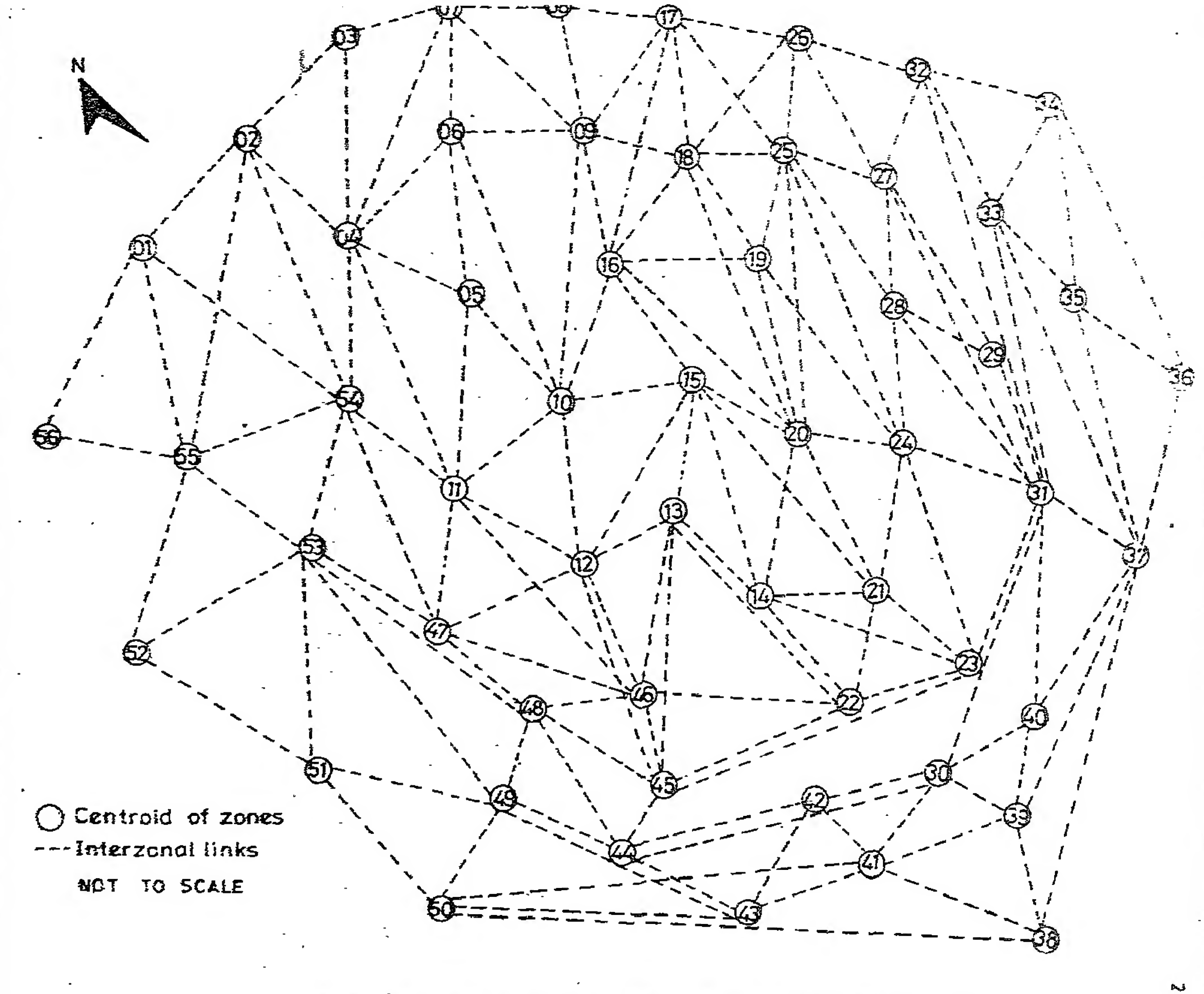


FIG-5.1 NETWORK OF CENTROIDS AND INTERZONAL LINKS

capacity should be specified also seperately for each mode. In heterogeneous traffic conditions, it is also difficult to specify a common measure of capacity. A great deal of planning is needed in estimating the link capacities. The city is having roads of only three standards, namely, dual carriage ways used by all modes, 2 to 3 lane wide roads having all modes of vehicles moving under stable flow conditions; and roads highly congested and having predominantly slow moving vehicles. Assigning the observed trip pattern along the shortest paths, the link flows are estimated. For each of the different road types the heavier link flows are identified and mode wise link capacities are estimated. Each link of a particular type may not have the same capacity level for a mode as the traffic compositions are varying on links. The observed link flows during peak periods also help in the final estimation of the mode wise link capacities. It may be worth mentioning that lot of judgement is involved in estimating these link capacities. The values as estimated are quite reasonable under the operating conditions. There is a great need to investigate this aspect in detail.

5.4.1.2 Selection of significant O-D pairs

The model involves a large magnitude of core capacity because of the requirements of the work vector, the coefficient matrices and resource vector inbuilt in the linear programming. The case study involves 312 unidirected links (m_1) and there are atmost 297 O-D pairs (n_1) that contribute the interzonal trips. This results in (m_1+n_1+1), that is, 610 inequality

constraints. The work vector dimension is $((M_1+M_2+2) * (M_1+M_2+2) + 3 * M_1 + 2M_2 + 4)$, where M_1 and M_2 are the inequality and equality constraints respectively. In this case M_2 is zeto as there are no equality constraints. It is observed that the available DEC-1090 system is not capable of handling the problems of this magnitude. It is noticed that the number of interzonal transfers for some of the O-D pairs are very insignificant. Further for some of the O-D pairs only one alternative path, that is, the shortest one is available. To reduce the magnitude of the problem without any significant impact on the results the following strategy is adopted.

(i) Given the trip generation by particular mode, the trips are distributed as per the distribution model to obtain t_{ij}.

(ii) t_{ij} matrix is inspected and those O-D pairs for which there is only one alternative path are seperated. O-D pairs having very insignificant interzonal transfers are also seperated. The remaining significant O-D pairs are used for further analysis.

(iii) The trip production and attraction values are modified by deleting the contribution of insignificant O-D pairs.

(iv) Given the revised trip generations, the combined trip distribution and assignment phase is carried for the remaining significant O-D pairs. Interzonal transfers and the link flows are estimated.

(v) The transfers associated with the deleted O-D pairs of step (ii) are assigned as per the shortest paths and combined with the values obtained from the model in step (iv).

As we are deleting O-D pairs contributing only insignificant transfers and assigning them by shortest path, the final results will not have any appreciable variation from those which could have been obtained by considering all the O-D pairs. For the case study, at most 100 significant O-D pairs are identified for some of the modes.

5.4.2 Gravity Submodel

To start with, the gravity submodel is to be calibrated so as to establish the relationship for the friction factor $f_{ij} = A\ \mathrm{EXP}(-\beta_0 c_{ij})$. The coefficients A and β_0 are to be calibrated by an iterative process until the simulated and observed travel time frequency distributions vary within $\pm$ 3 percent. Using the observed trip generation matrix and some assumed values of A and β_0 (A = 1.00 and β_0 =0.30), the trip matrix, t_{ij}, is evolved and frequencies for each travel time group are determined. The value of friction factor for each of the travel time group is revised as follows.

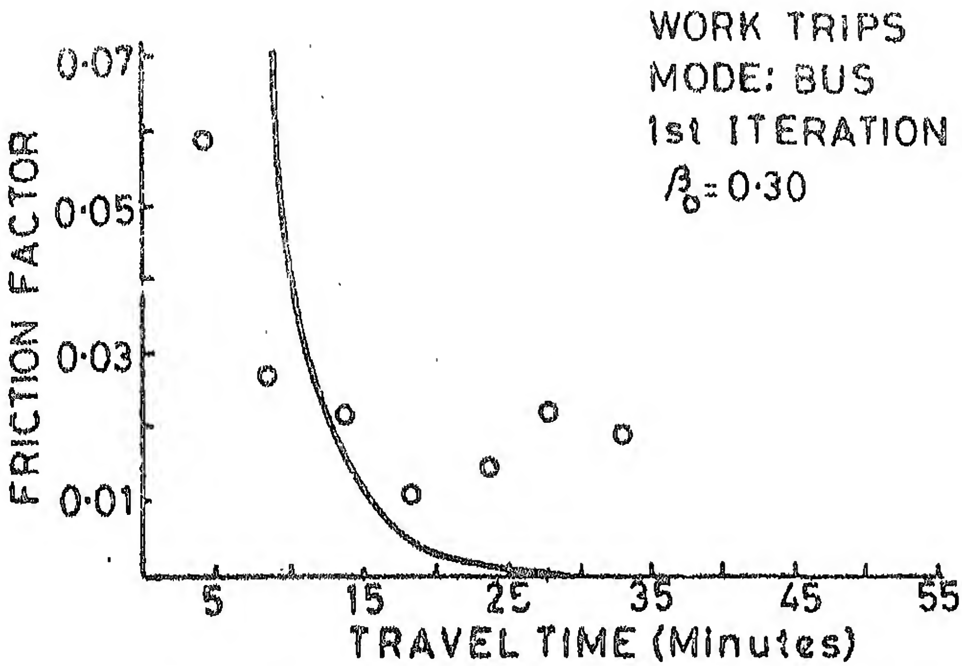
For a particular travel time group (i) the friction factor $f_i' = f_i \dfrac{OD_i}{GM_i}$ as given in Eqn.(5.19). A new friction factor of the form as in Eqn.(5.11) is to be fitted to all the revised friction factor values. This relationship is established by an iterative process, using the DFP method

with the penalty parameter, r, to find the search direction $\bar{S}$ and cubic interpolation method to find the appropriate step length α. A number of iterations are to be made for estimating the calibration coefficients A and β_0. The friction factor-travel time for some of the iterations in case of work trips by transit mode are shown in Figs. 5.2 to 5.7. A comparison of the observed and simulated travel time frequency distribution for some of the iterations are also shown in Figs. 5.8 to 5.11. It is observed that with each iteration the simulated values come closer to the observed ones. In case of final iteration (Fig.5.11), the observed and simulated frequencies are within the tolerance of $\pm$ 3 percent. It is noted that the value of A is always unity. The final calibrated values of β_0 are given in Table 5.1 for each of the modes, for both work and education trips. The number of iterations varied between 6 and 18 and the value of R^2 is very high. The procedure of constrained optimization as adopted in this study is definitely better than the other usually adopted procedures (sub-section 5.3.1.2). The calibrated values of β_0 are given as input to the combined distribution and assignment model.

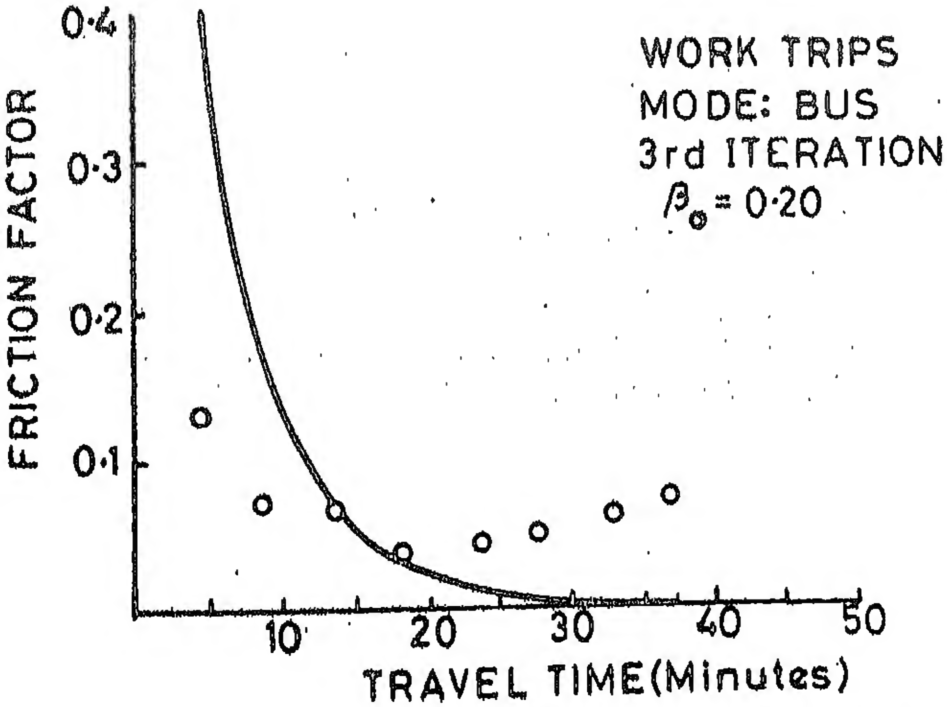
The next task is to choose the entropy constraint E_0. The choice of E_0 is synonymous to the calibration of the gravity submodel. Starting with $\beta = \beta_0$ (as calibrated above), the t_{ij} matrix is of the form

$$t_{ij} = O_i \; D_j \; \text{EXP} \; (-\beta \; c_{ij}) \qquad .. \; (5.68)$$

The constraints O_i and D_j are adjusted till the marginal totals are obtained (sub-section 5.3.1.3). This trip matrix has an



FIG·5·2 FRICTION FACTOR-TRAVEL TIME RELATION



FIG·5·3 FRICTION FACTOR-TRAVEL TIME RELATION

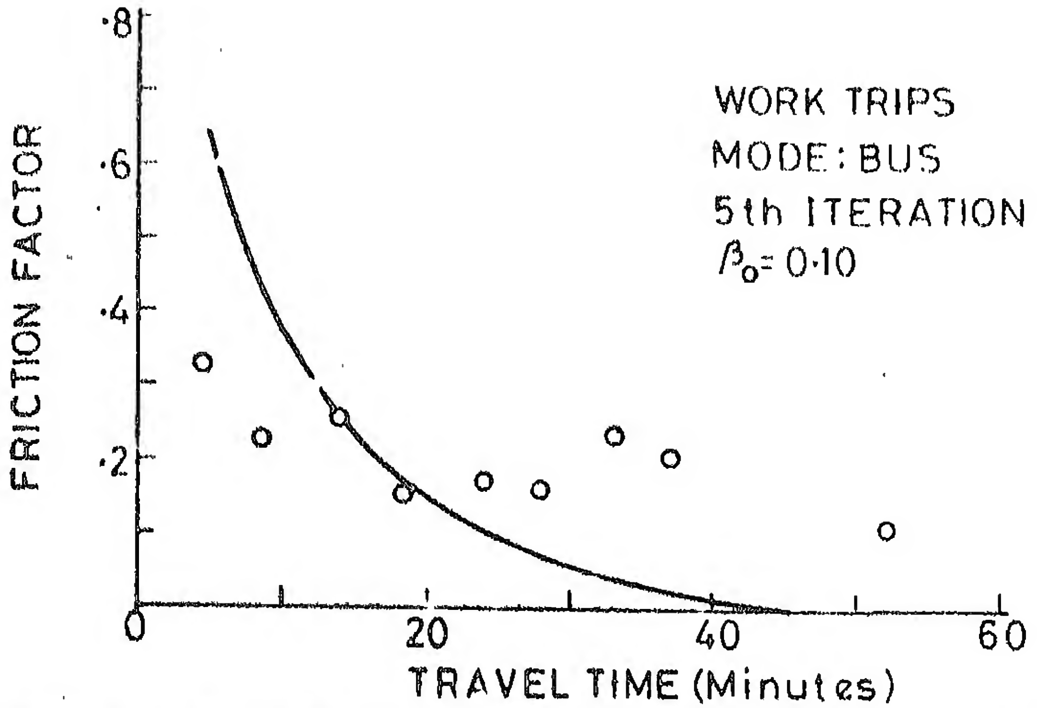


FIG. 5·4 FRICTION FACTOR – TRAVEL TIME RELATION

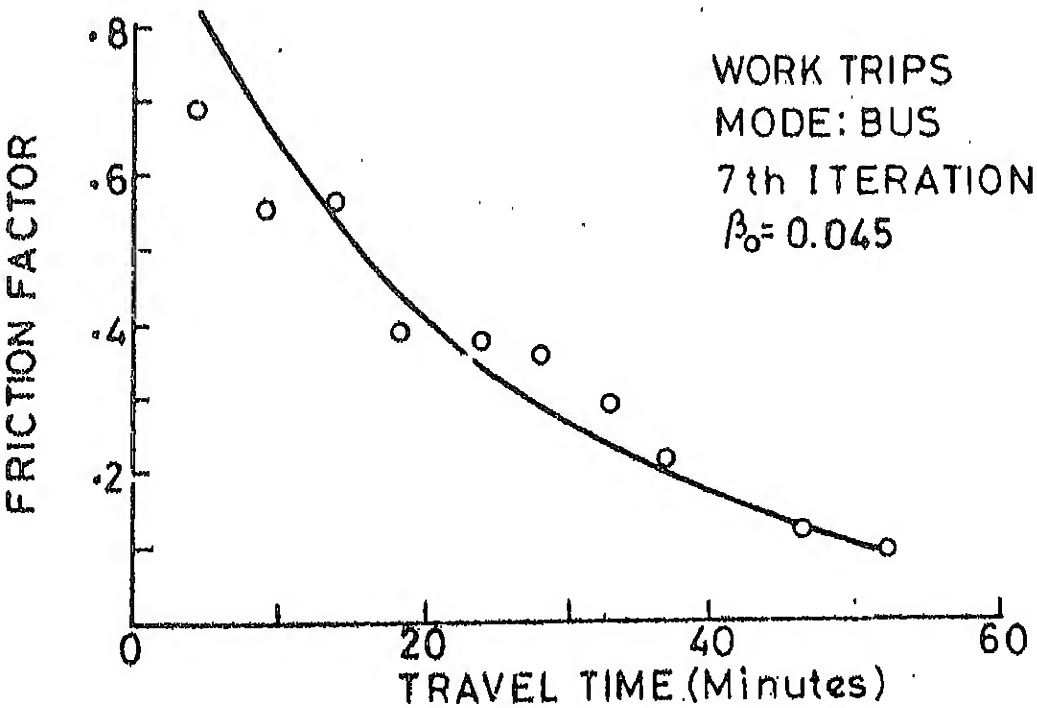
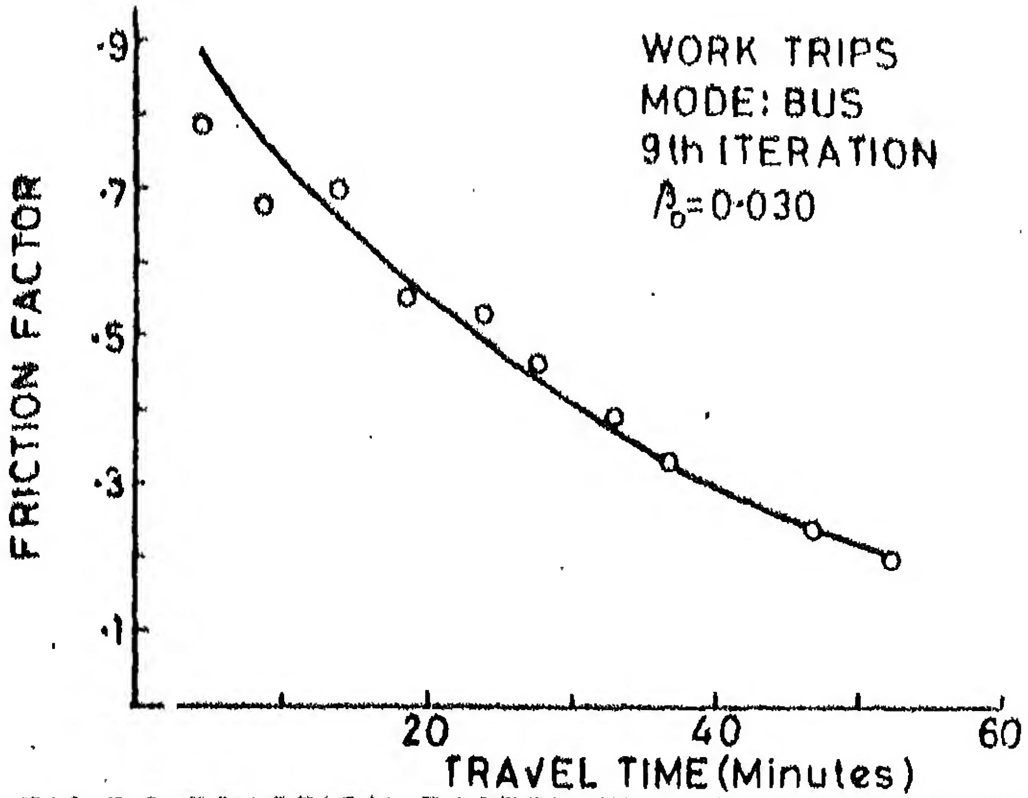
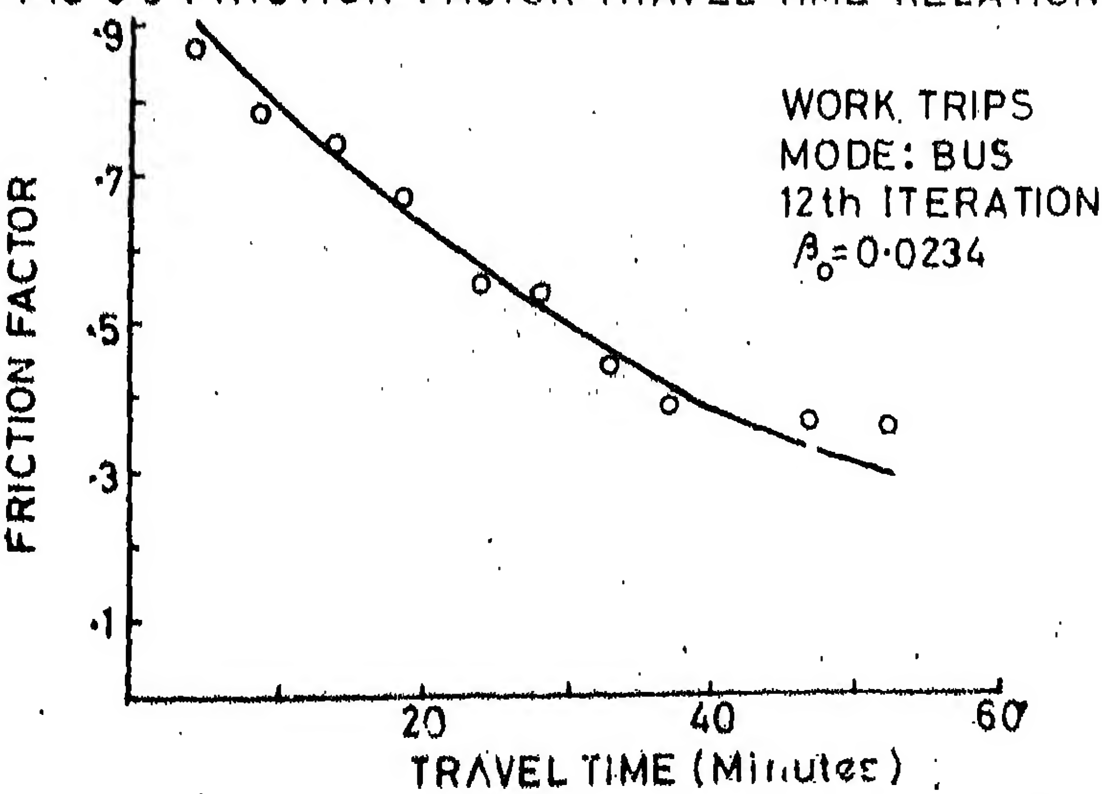


FIG. 5·5 FRICTION FACTOR – TRAVEL TIME RELATION



FIG·5·6 FRICTION FACTOR-TRAVEL TIME RELATION



FIG·5·7 FRICTION FACTOR-TRAVEL TIME RELATION

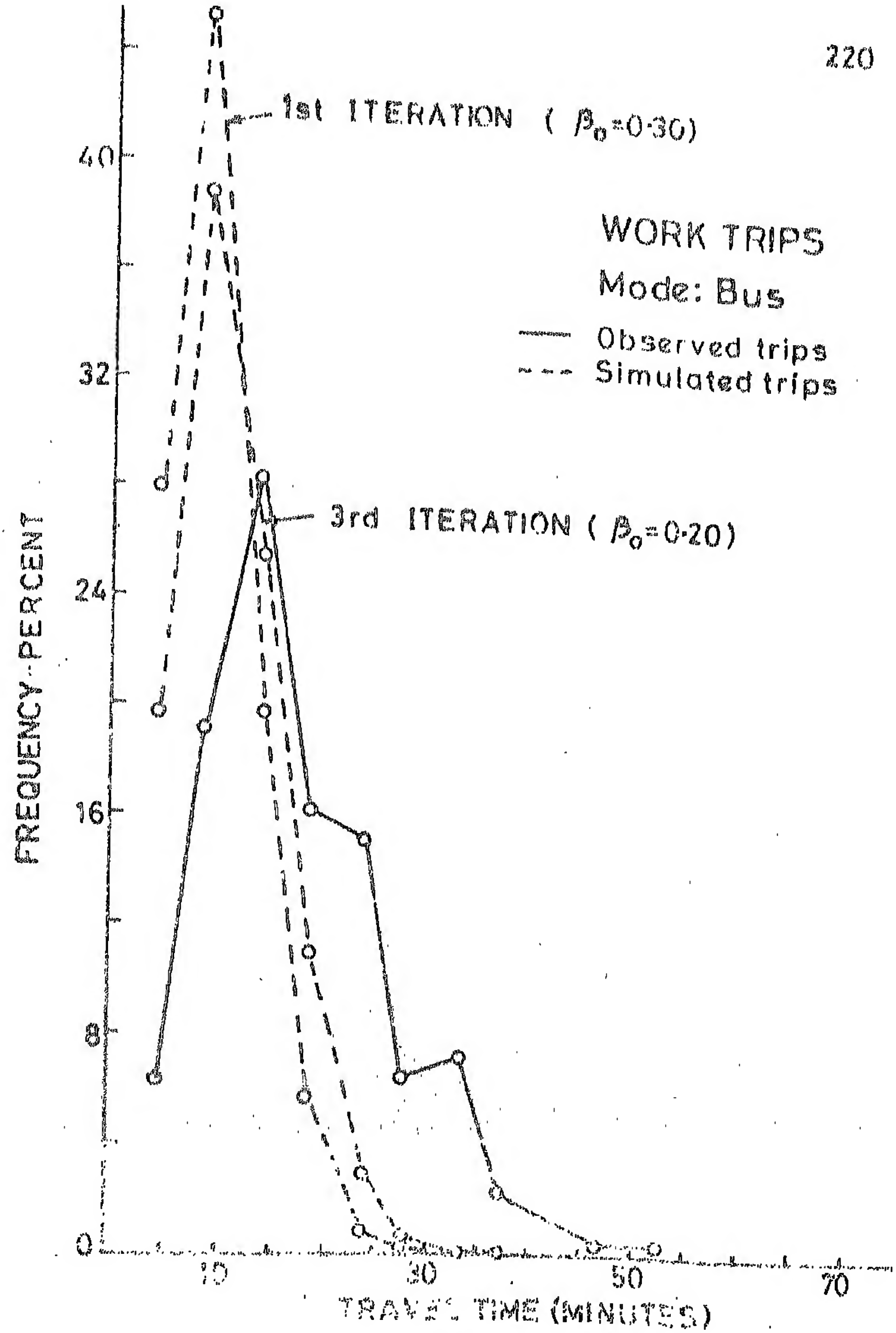


FIG-3·8 TRAVEL TIME FREQUENCY DISTRIBUTION
(1st and 3rd Iteration)

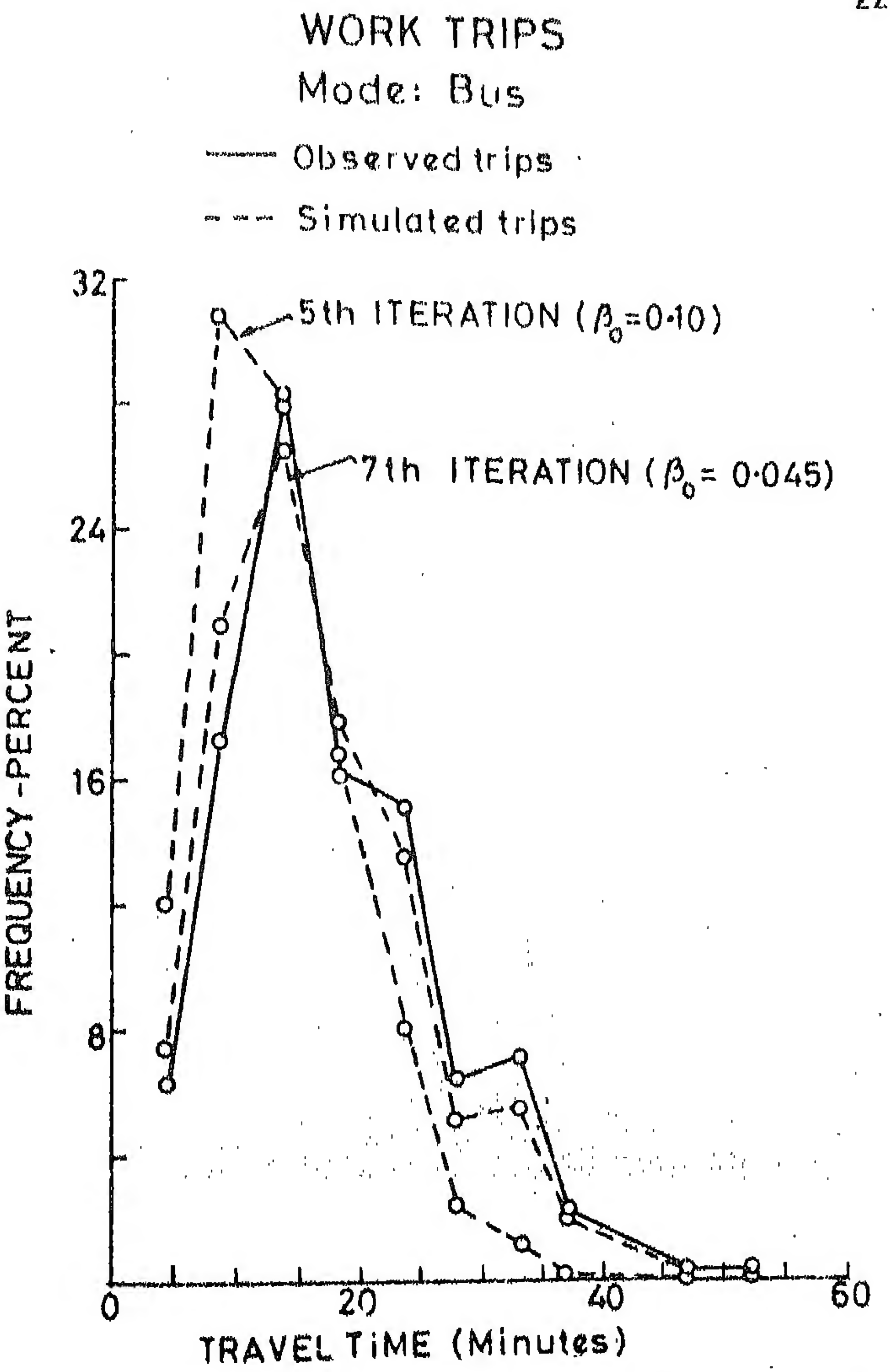


FIG.5.9 TRAVEL TIME FREQUENCY DISTRIBUTION
(5th and 7th Iteration)

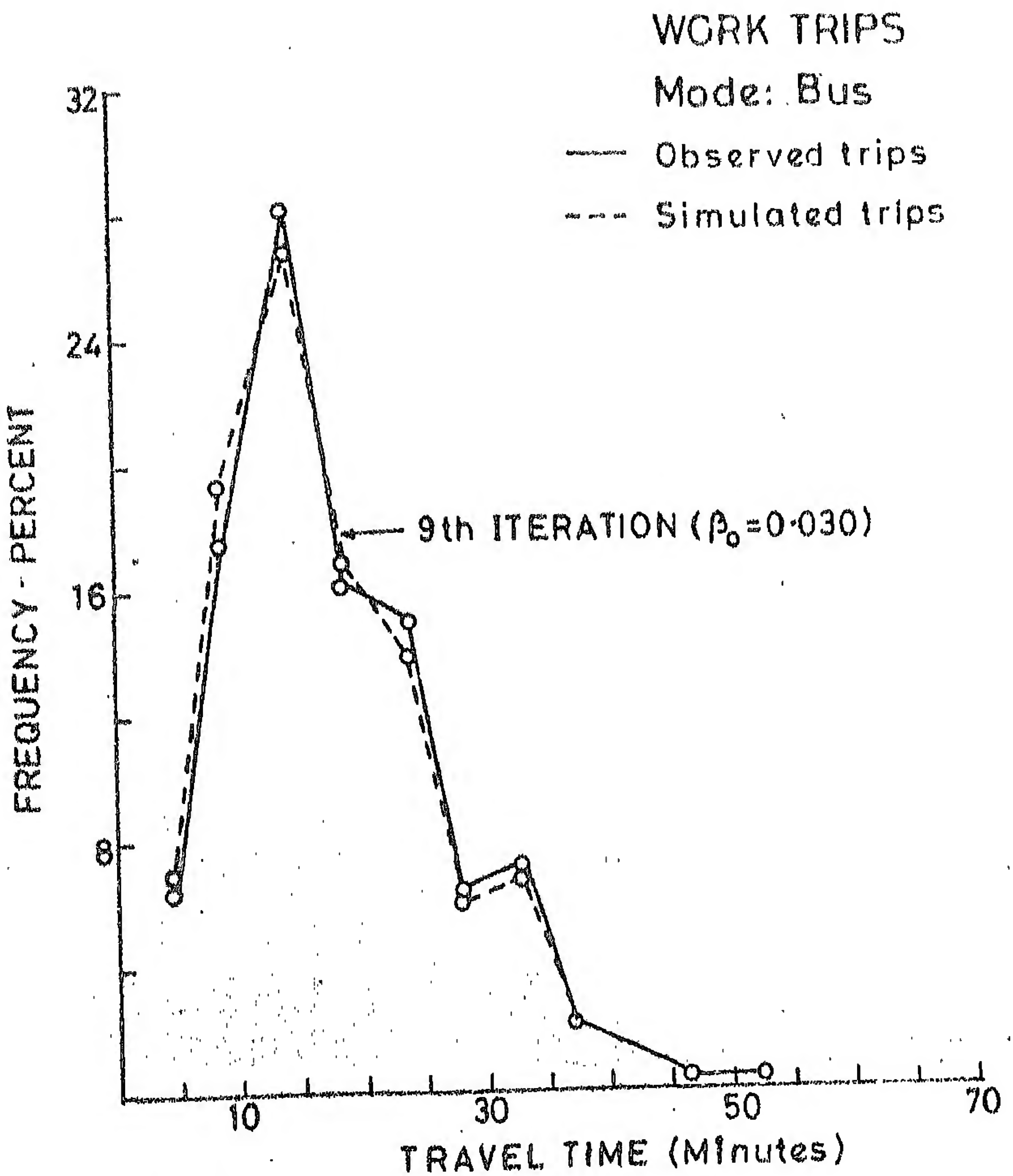
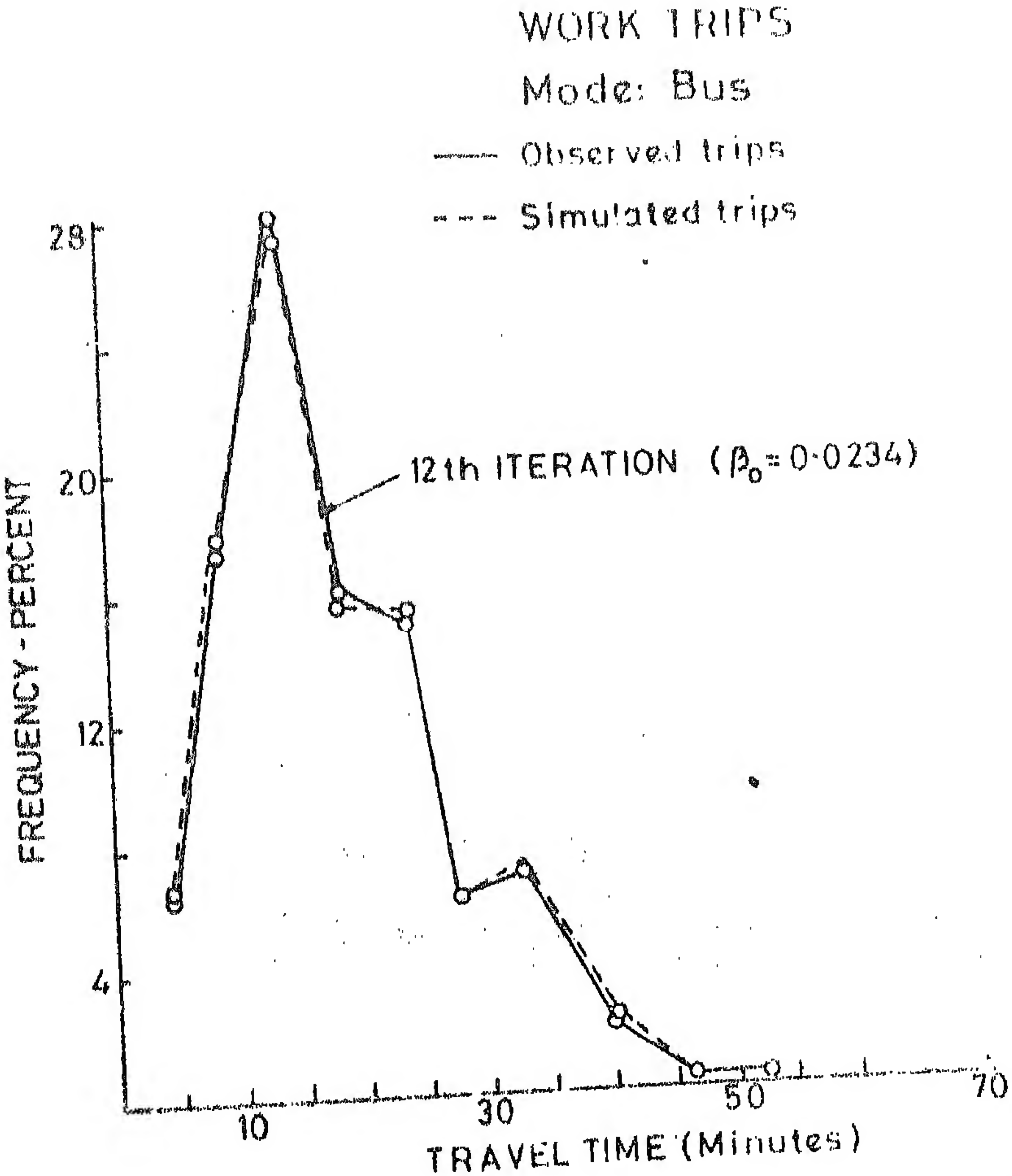


FIG.5·10 TRAVEL TIME FREQUENCY DISTRIBUTION (9th Iteration)



FIG·5·11 TRAVEL TIME FREQUENCY DISTRIBUTION
(12 th Iteration)

TABLE 5.1 : FINAL CALIBRATED COEFFICIENT(β_0) FOR THE RELATIONSHIP f_{ij} = A EXP $(-\beta_0 c_{ij})$

Trip purpose	Mode	Final β_0	R^2	No. of iterations
Work trips	Walk	0.0358	0.98	12
	Bicycle	0.0560	0.97	15
	Motorcycle/ scooter	0.1079	0.99	18
	Car	0.1216	0.96	8
	Bus	0.0234	0.98	12
	Cycle rikshaw	0.0700	0.97	15
Education trips	Walk	0.0780	0.97	6
	Bicycle	0.0782	0.96	6
	Motorcycle/ scooter	0.3744	0.96	9
	Bus	0.1011	0.98	9
	Cycle rikshaw	0.1251	0.97	12

accessibility $E_0 = -E(t_{ij}(\beta))$. The system demands that the accessibility E should be at least what it is at present (E_0). The functional value $-E(t_{ij}(\beta))$ is calculated in the following way.

Taking the first approximation of t_{ij} as T_{ij}, which is taken as,

$$T_{ij} = A \text{ EXP } (-\beta \, c_{ij}) \qquad .. (5.69)$$

This is revised through an iterative procedure (sub-section 5.3.1.3) which stops when row sums and column sums are close to the marginal totals such that

$$p_{max} \, a_{max} \text{ and } \frac{1}{p_{min} a_{min}} < 1 + \frac{1}{6T}$$

where,

$$p_{max} = \max_i \frac{p_i}{T_{i*}}$$

$$a_{max} = \max_j \frac{a_j}{T_{*j}}$$

$$p_{min} = \min_i \frac{p_i}{T_{i*}}$$

$$a_{min} = \min_j \frac{a_j}{T_{*j}}$$

$$T_{i*} = \sum_j T_{ij}$$

$$T_{*j} = \sum_i T_{ij}$$

p_i = productions in zone i

a_j = attractions to zone j

$$T = \sum_i \sum_j T_{ij} \qquad .. (5.70)$$

After having obtained the solution $T_{ij}(\beta)$, the corresponding entropy $- E(T_{ij}(\beta))$ is computed.

Using the iterative procedure (sub-section 5.3.1.3) the solution of the trip distribution model for each of the mode for work and education trips are obtained. The maximum and the planned entropy values are given in Table 5.2. In each of the set the planned accessibility indices, E, are higher than the existing values E_0. Each solution involved a number of iterations ranging between 15 to 32. It is observed that for education trips more number of iterations are performed than those for the work trips. This is due to less number of education trips which makes the convergence slow.

5.4.3 Combined Distribution and Assignment Model

5.4.3.1 Selection of significant node pairs

The combined model involves the interaction of the gravity submodel and assignment submodel. Due to large requirement of the work vector, it is decided to ignore some of the insignificant node pairs while solving for the combined distribution and assignment model. The t_{ij} matrices of different modes both for work and education trips as obtained from the solution of the trip distribution model (sub-section 5.3.1.3) are inspected and those pairs which have very little interzonal flows or for which there is only one feasible path are eliminated. Taking only the remaining significant node pairs the trip generation values are updated to be provided as input for the combined model. Further only the trips performed by car,

TABLE 5.2 : MAXIMUM AND PLANNED ENTROPY VALUES

Trip purpose	Mode	Maximum entropy (E_{max})	Planned entropy (E)	No. of iterations
Work trips	Walk	6.701	5.725	24
	Bicycle	6.818	6.296	15
	Motorcycle/ scooter	6.499	5.445	16
	Car	5.054	3.618	25
	Bus	6.387	4.982	24
	Cycle rikshaw	6.472	4.978	20
Education trips	Walk	6.548	4.986	30
	Bicycle	6.355	5.332	18
	Motorcycle/ scooter	4.788	3.052	28
	Bus	5.892	4.749	26
	Cycle rikshaw	6.401	4.784	32

bus and motorcycle/scooter are taken for model solution. Atmost 100 significant node pairs are selected for some of the modes. In some cases the significant node pairs are even less.

The computations of trip distribution submodel have so far used the shortest path for the generalized cost matrix c_{ij}. The solution of the programming model is aimed through successive iterations of trip distribution and traffic assignment models as per the algorithm described in sub-section 5.3.2.1. Given the list of acceptable alternative paths for each of the concerned O-D pairs, the assignment submodel is run to evolve new costs c_{ij}. These in turn find their way to replace the earlier values. The gravity submodel is run again which provides a new distribution matrix to feed the assignment submodel. The iterative process gets terminated when a stable solution is obtained.

Let ($\bar{x}_q^{ij}$, $\bar{\lambda}_k$) and ($\bar{\pi}_p$, $\bar{\alpha}_{ij}$, $\bar{\delta}$) (sub-section 5.3.2.1) be the solution to the assignment problem and its dual respectively for a particular case. The value of the dual variable $\bar{\pi}_p$ has the link capacities y_p , as its coefficients. The dual cost coefficients $\bar{\alpha}_{ij}$ significantly affect the objective function. The initial values of α_{ij}^0 is chosen according to the shortest path algorithm and δ^0 can be set to ∞. To minimize the number of variables only two alternative paths are considered between each node pair. The shortest path is always considered and the second is the next best to the shortest. When solution to the linear programming problem is obtained, it is checked if any more paths should enter the problem by

calculating the reduced cost for all the variables. The reduced cost for a variable is

$$\sum_p c_p \, a^{ij}_{pq} + \sum_p \pi_p \, a^{ij}_{pq} - \alpha_{ij} \quad .. (5.71)$$

If for some path $\sum_p (c_p + \bar{\pi}_p) \, a^{ij}_{pq} < \bar{\alpha}_{ij}$, then this path enters the set of alternative paths. The paths which is the shortest between i and j when the link costs are $(c_p + \bar{\pi}_p)$ is the one with the smallest value of $\sum_p (c_p + \bar{\pi}_p) \, a^{ij}_{pq}$. The algorithm is terminated at a particular iteration, say k, when

$$\sum_i \sum_j \alpha^{k-1}_{ij} \, t^k_{ij} > \delta^{k-1} \quad .. (5.72)$$

For the case study the iteration process of trip distribuion and assignment is generally terminated in only 2 to 3 iterations. The convergence is so fast because not many alternative paths are available for the given road network. The CPU time required for each iteration is about 7 minutes, when 100 significant node pairs are considered. The algorithm is computationally quite efficient and is capable for application even when more alternative paths could be made available. The assignment is done by system optimizing principle and the value of the objective function along with the number of alternative paths entering the solution for each of the set are given in Table 5.3.

Trial runs indicated that the values of link capacities significantly affect the number of alternative paths entering the final solution. Due to various constraints of time and resources, the detailed investigations on the sensitivity of the link capacities could not be done. The model output

TABLE 5.3 : OBJECTIVE FUNCTION FOR DIFFERENT MODES

Trip purpose	Mode	Objective function value	No. of alternative paths entering the solution
Work trips	Motorcycle/ scooter	187488	103
	Car	36786	50
	Bus	182125	107
Education trips	Motorcycle/ scooter	10740	23
	Bus	166237	105

indicates that the final t_{ij} values are considerably changed than those obtained in the initial run of the distribution submodel. The accessibility, that is, the entropy E as explained in sub-section 5.3.1.1 is computed for the model output of the t_{ij} matrix. This is compared with the maximum value of the entropy, E_{max}, and that of E_0 for the initial trip matrix. The values of the entropy for different sets are given in Table 5.4. The planned entropy E is significantly higher than the existing entropy E_0 for all the cases.

5.4.3.2 Analysis of Model output

A model is of relevance only when it is validated. For the case study of Kanpur, the available household data provides the trip distribution matrices for different modes of travel, but no information is available about the paths chosen by the trip makers. It is thus possible only to have the validation of the trip distribution aspect which is interacting with the traffic assignment phase in the model.

The observed and simulated frequencies of trips for each of the travel time groups are shown in Figs. 5.12 to 5.15 for different trip types by each mode. The comparisons indicate that the simulated values are highly close to the observed ones. A comparison of the observed and simulated interzonal transfer values, i.e., t_{ij} of work trips for different modes are shown in Figs. 5.16 to 5.18. Comparison of the work trips for interzonal transfers by bus and car shown in Figs. 5.16 and 5.17 indicate that the t_{ij} values as obtained from the model are within ± 10 percent of the observed ones. Further the

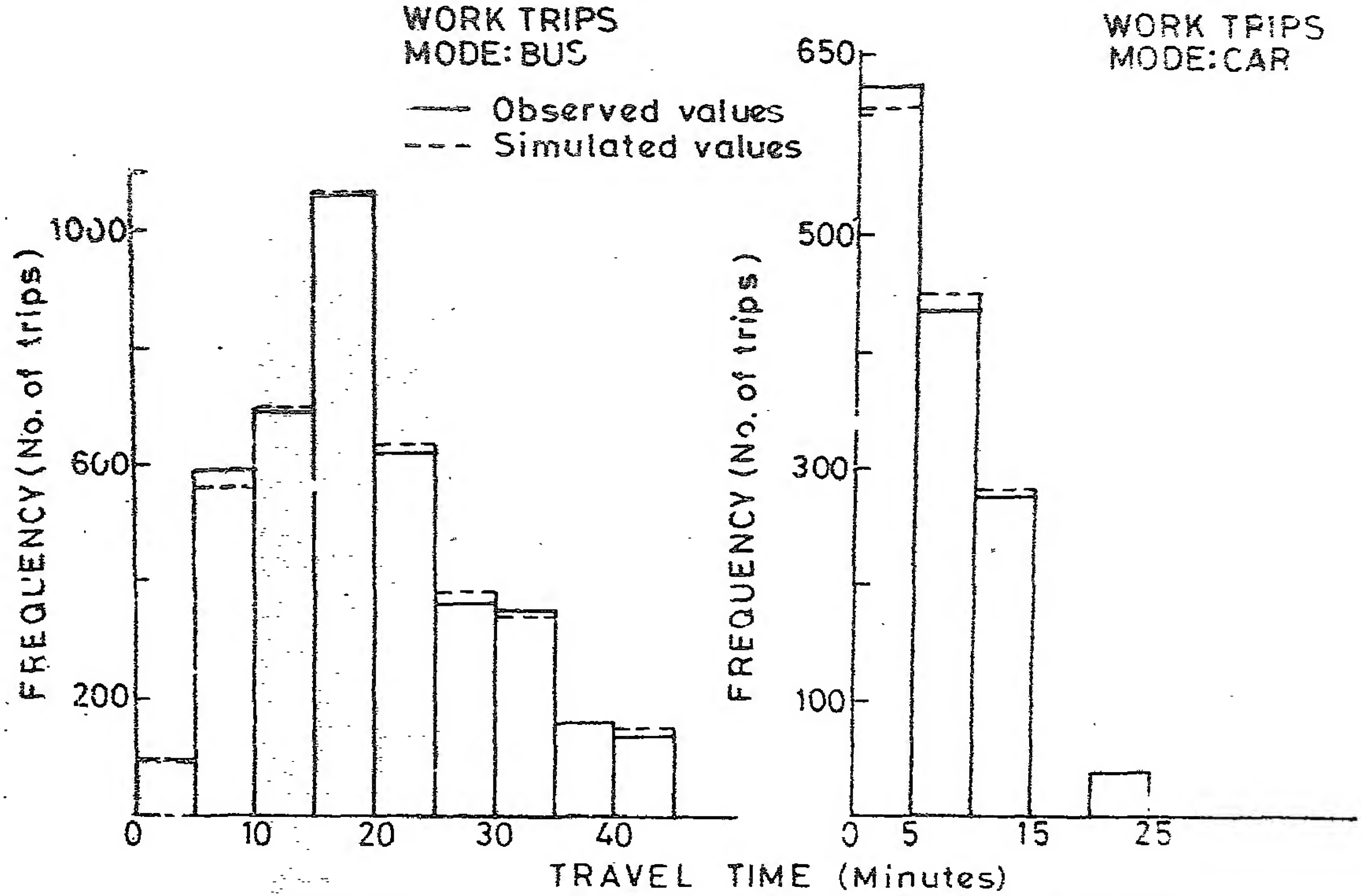


FIG-5-12

FIG-5-13

FREQUENCY DISTRIBUTION OF DIFFERENT TRAVE TIMES

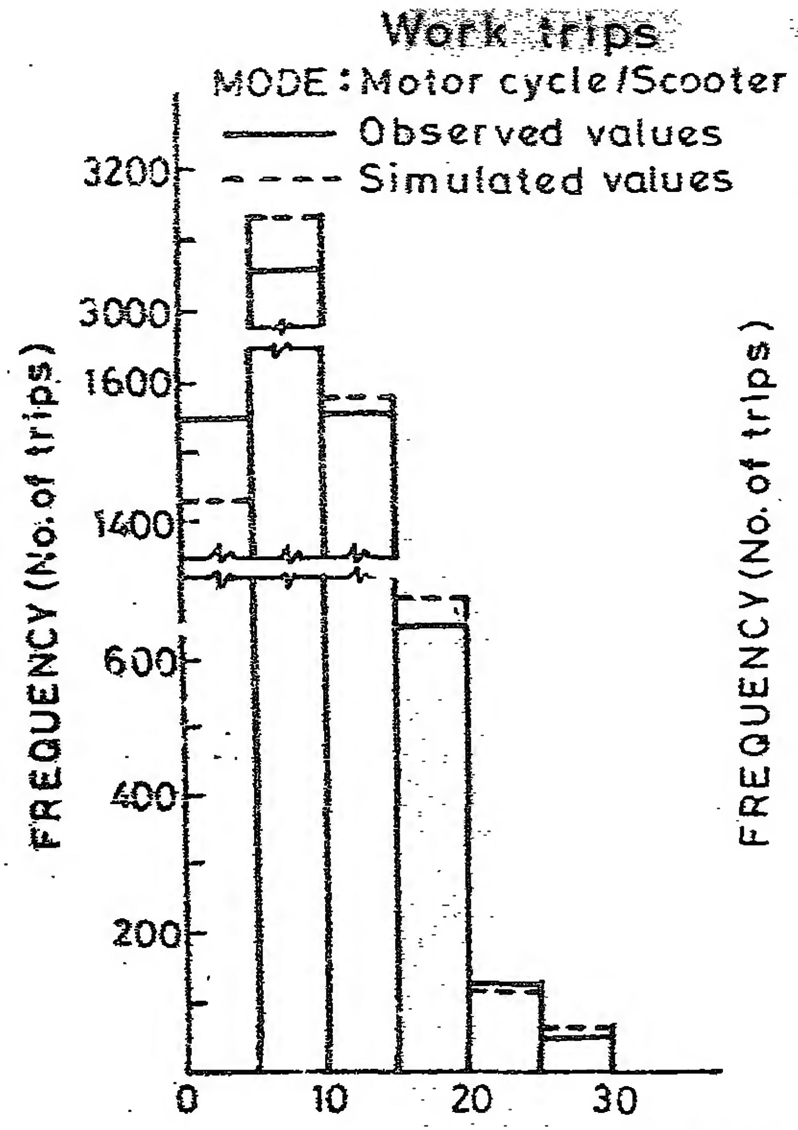


FIG. 5·14

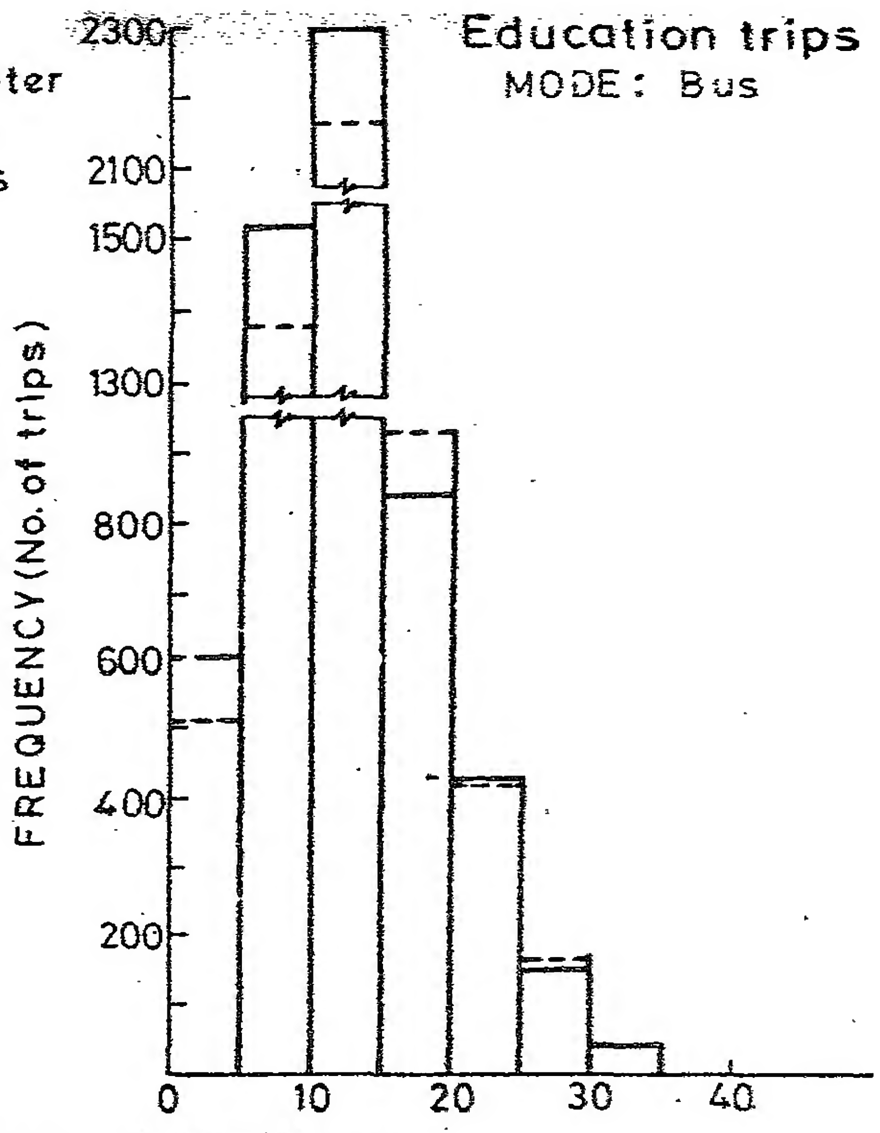


FIG. 5·15

FREQUENCY DISTRIBUTION OF DIFFERENT TRAVEL TIMES

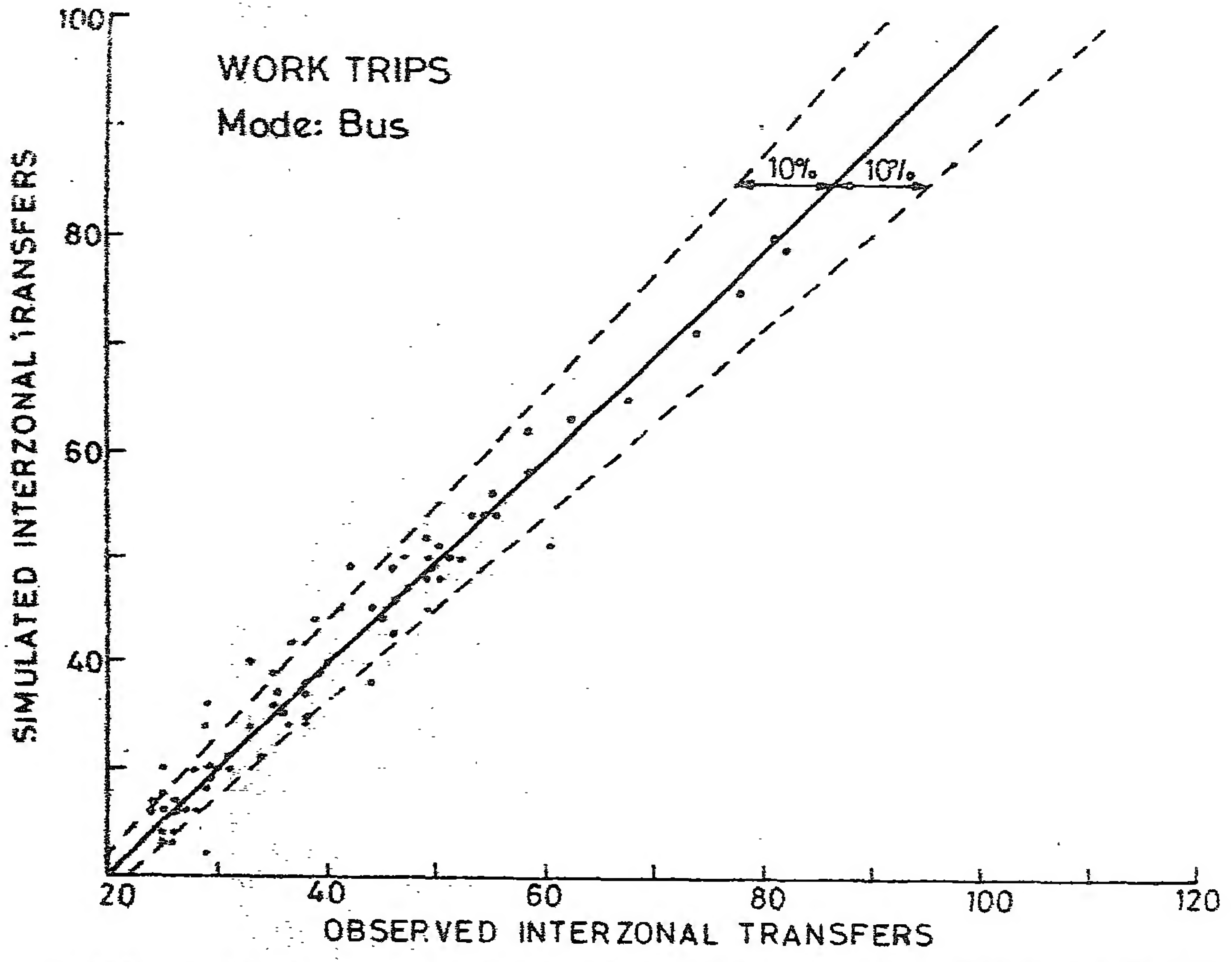


FIG-5-16 COMPARISON OF OBSERVED AND SIMULATED INTERZONAL TRANSFERS

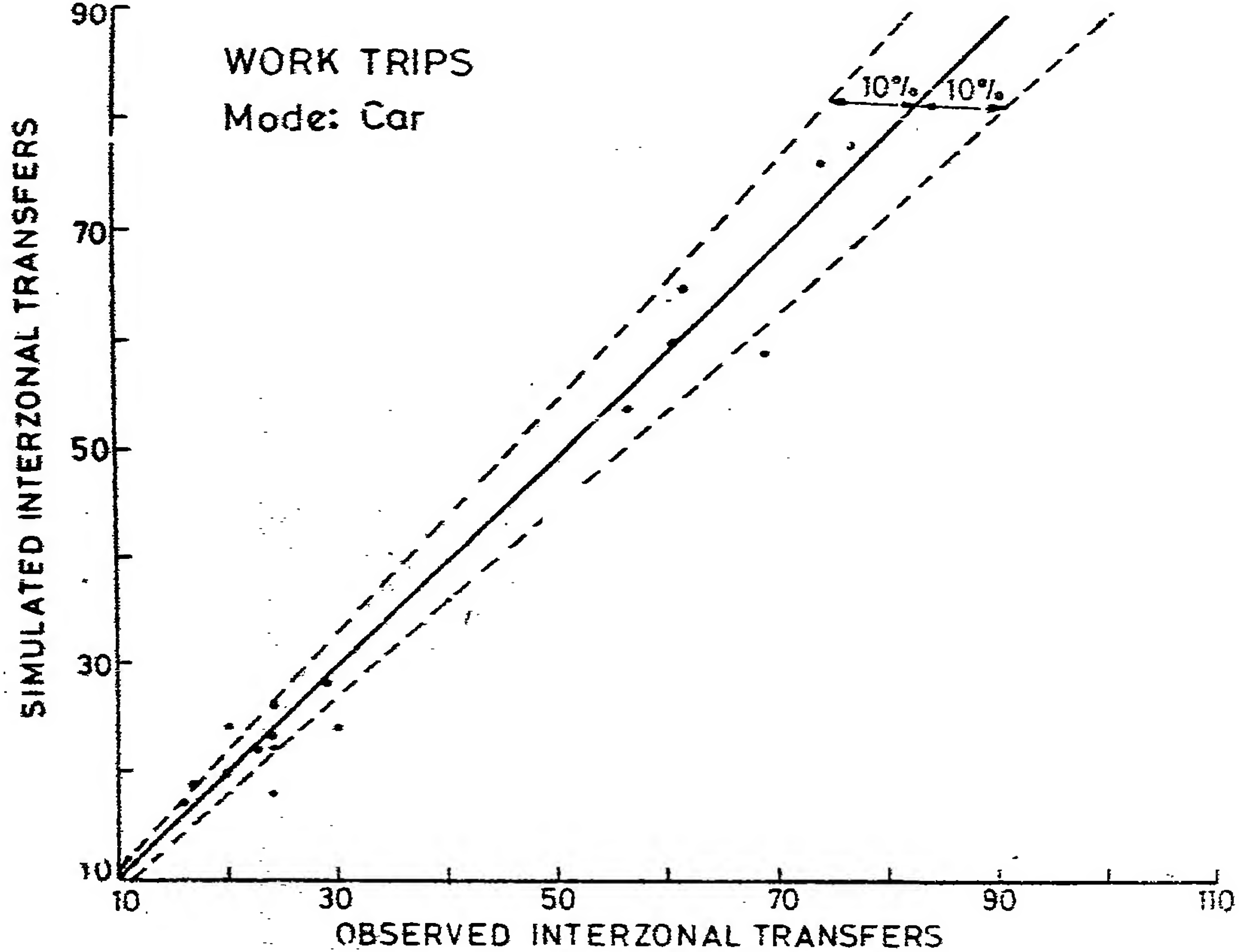


FIG.5.17 COMPARISON OF OBSERVED AND SIMULATED INTERZONAL TRANSFERS

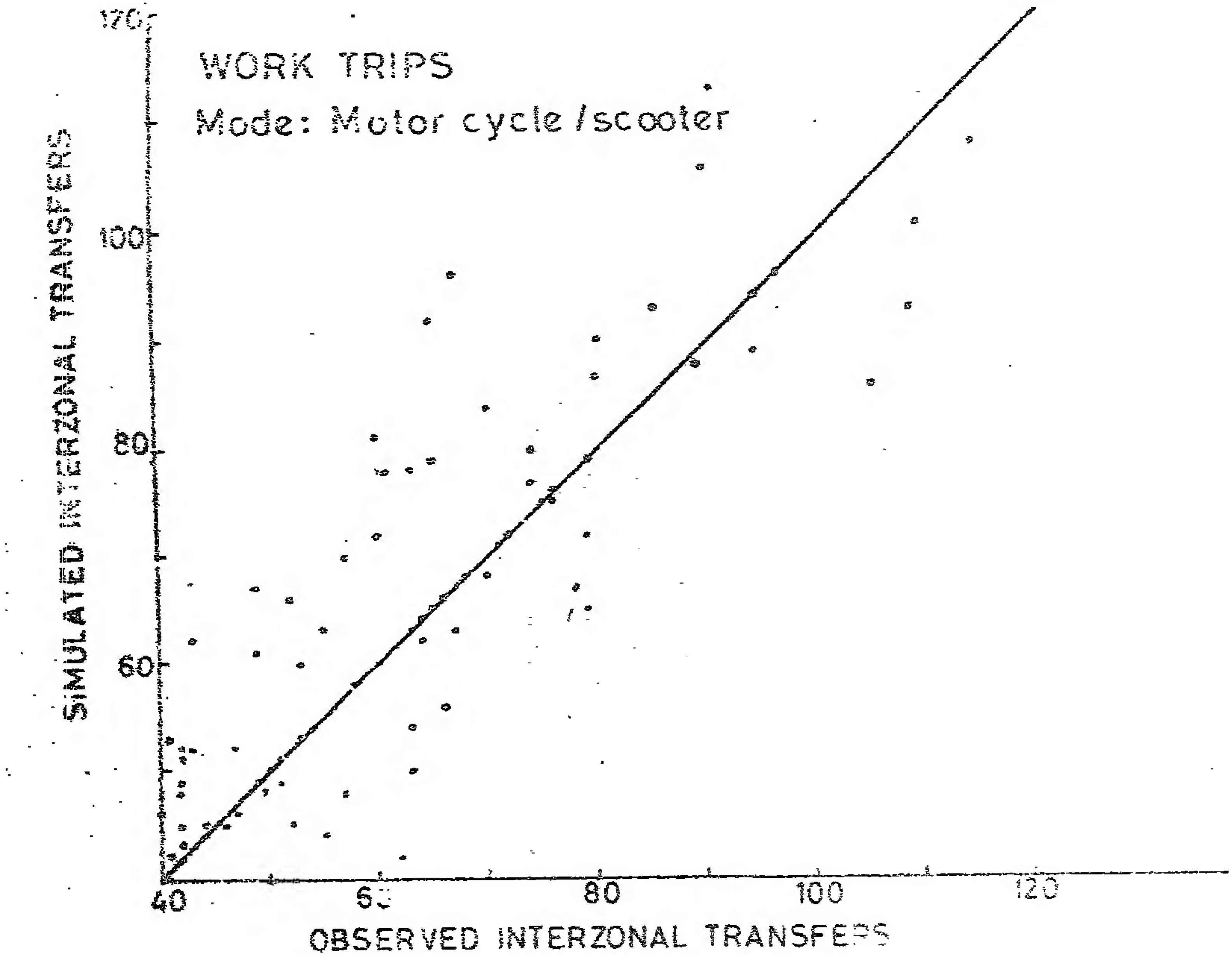


FIG. 5.18 COMPARISON OF OBSERVED AND SIMULATED INTERZONAL TRANSFERS

values are well distributed on both sides of the 45 degree line indicating a good fit. In case of work trips by motorcycle/scooter, the variation is slightly higher, but equally distributed on each side of the 45 degree line. This comparison indicates that not only the simulated travel time frequencies are close to the observed ones, but each value of the t_{ij} matrix also compares well. A sufficiently high value of the entropy further enhances the capability of the model.

The assigned flows on each of the alternative paths x_q^{ij}, for the significant O-D pairs are obtained from the model for each of the sets. The t_{ij} values for the insignificant node pairs are assigned by the shortest path as explained earlier. From the complete assignment, the flows on each link are determined. Fig. 5.19 shows the flow on different links for work trips by motorcycle/scooter as finally obtained, along with the values of the flows, when all interzonal transfers are assigned along the shortest paths. The comparison of the two link flows indicates that even for the city of Kanpur, where not many alternative paths are available, flows on some of the links do get affected while optimizing the system costs. The detailed analysis could not be carried out due to lack of information on the paths chosen by the trip makers.

The results indicate that the algorithms for the combined distribution and assignment models are computationally quite efficient. Further the strategy adopted for selecting the significant O-D pairs and then having a combined distribution and assignment for these pairs are quite efficient for the large size networks.

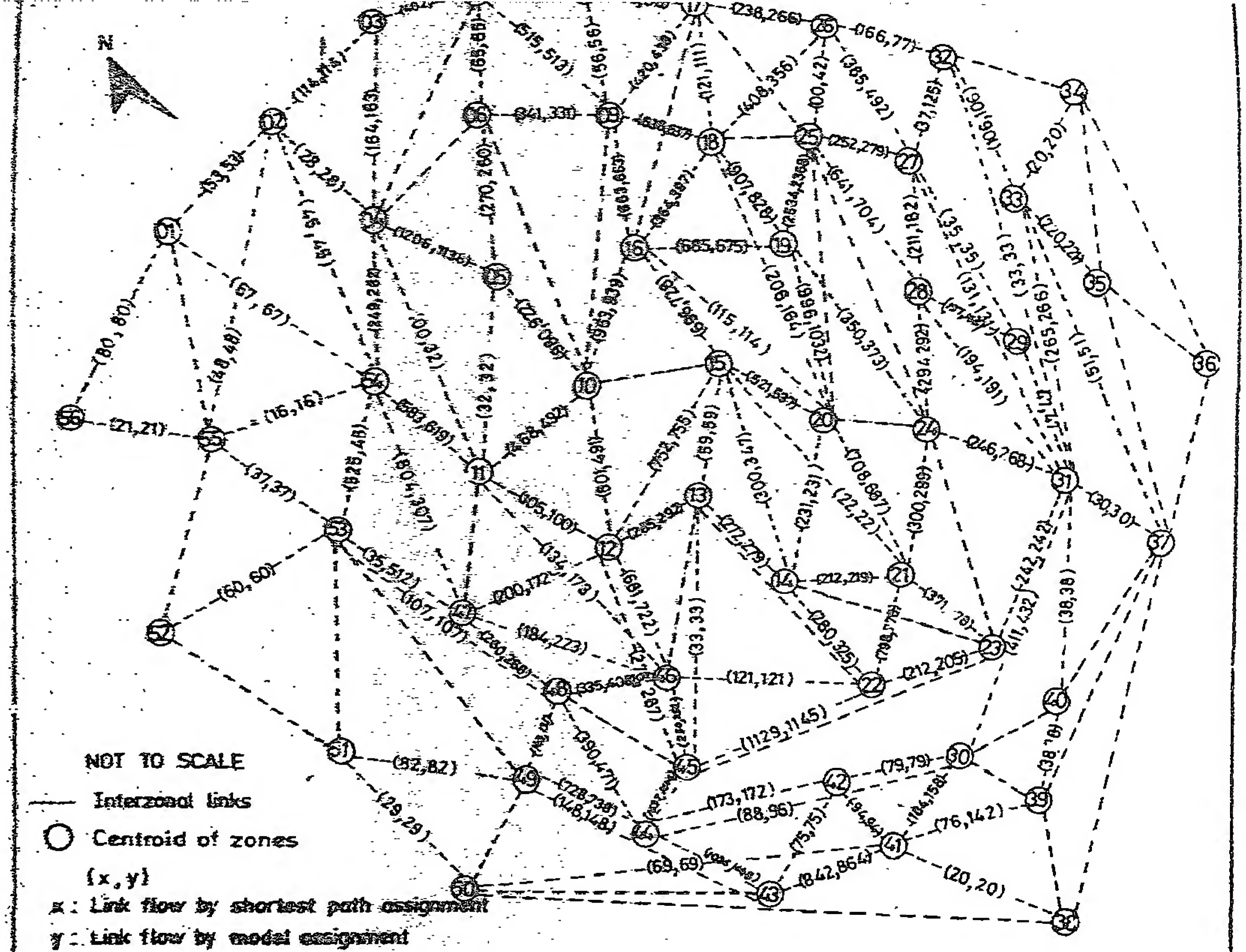


FIG-5-19 FLOWS ALONG THE INTERZONAL LINKS OF THE NETWORK

6 SUMMARY AND CONCLUSIONS

6.1 SUMMARY

Process of urbanization in developing countries is not simply a repetition of what occurred in the developed world, rather it is occurring much faster and under different economic and technological conditions. Large size developing cities generally suffer more from traffic congestion, road accidents, environmental pollution, overcrowding of public transport, and poor conditions for pedestrians.

In the last three decades planning of urban transport has changed dramatically and a large number of models have been developed for various stages of the travel forecasting process. These models depend on a high level of car ownership and a relatively homogeneous distribution of land use at low density. The cities in developing countries have lower income level, multiple modes of travel, heterogeneous distribution of land use at high density and poor transport facilities. The factors that generate trips and which determine mode choice among the low income groups are not well understood and cannot be deduced from experience in developed countries. A study of Indian social structure in relation to travel habits is necessary to develop trip generation models.

The aim of this study is to develop system of models appropriate for Indian urban structure, which given a knowledge

of land use, household characteristics, and available transport facilities could estimate accurately the intricate movements of traffic on the system by mode. The models evolved in the study include : trip production models; mode choice decision models based on non-compensatory choice processes; simulation of modal split and trip distribution as an interconnected decision process; and programming model for combined trip distribution and assignment. These models could be used to show what would happen if transport facilities are changed and if over time, the land use and household characteristics are altered. The efficient methodologies for travel demand models can be evolved only when carefully programmed basic data related to the existing situation of the system are available. The models evolved are calibrated and validated for the city of Kanpur, the premier industrial metropolis of Northern India. The city is divided into 56 traffic zones and the models are calibrated for the base year household travel survey data.

Trip Production Models.

Three trip production models, namely, zonal regression category analysis, and household regression models are constructed based on the household survey data. For the selection of best set of explanatory variables in the zonal and household regression models, forward stepwise regression method is used. The explanatory variables considered for zonal regression are: number of persons in the zone, total income level for the zone, number of vehicles of different types in a zone, and zonal trip length. Zonal trip production relationships are obtained for the work and

education trips by modes. Considering the type of household data available and with the aim to conserve homogeneity within each type of household, category analysis is used for productions of work and education trips. The household is classified in three different ways such as family size-income level, family size-vehicle ownership, and income level-vehicle ownership combinations. In household regression approach the analysis is directly applied to the basic household data, and it operates on the whole of the variability which exists in the data at that level. This approach is a combination of better features of zonal regression and category analysis. The choice of explanatory variables are guided by the results of zonal regression and household classifications. The explanatory variables that have been considered are: family size, income level, vehicle ownership, and trip length. These variables are further stratified into 21 subgroups. Because of correlation between the variables, number of explanatory variables that are found to enter the regression equations range between seven and fifteen.

Mode Choice Decision Models

Mode choice decision models of non-compensatory structures are made on the basis of attribute by attribute comparison of various alternatives. The levels of satisfaction associated with each of the attributes perceived by the person in the decision process are evaluated. Based on the concept of decision theory maximin, maximax, and predominant-attribute models are formulated. The choice of attributes involved in the decision making process is of paramount importance. In a typical metropolitan city of

India having number of alternative modes, the decision maker has to access his/her own socio-economic characteristics in relation to the characteristics of available alternative modes. The attributes considered in this study are: family size, income level, vehicle ownership, and trip length.

The maximin/maximax strategy calls for the identification of least/highest satisfactory alternative of each attribute and the selection of the alternative which yields the highest minimum/maximum level of satisfaction. Predominant-attribute model calls for identification of the most predominant-attribute according to the generic importance. Alternative having the highest level of satisfaction for the predominant attribute is selected. The results obtained from maximin and maximax models compare well with the observed data, but the results obtained from the predominant-attribute model significantly vary from the observed values. As such the most predominant-attribute model as formulated does not appear to be suitable for the city of Kanpur. This may be attributed to the fact that the utility value for each attribute is taken as the proportion of trips for that alternative, or it may be that the mode choice decision process is not affected only by the most predominant attribute. It is desirable to test this model for some other cities also to identify the appropriate causesfor differences.

Simulation of Modal Split and Trip Distribution

Different designs for disaggregate models have been suggested to solve complex interaction processes, but the

implementation has been highly hampered by the methodological difficulties. Simulation of the complex decision process at the individual level is attempted in this study. The model simulates modal split and trip distribution as an interconnected decision process. The model essentially consists of status group, mode choice, and distribution submodels. Individuals with similar constellations of activity constraints are grouped into status groups. The individual decision process which is stochastically determined from the decision profile of his/her status group, identifies the mode choice and trip destination. The decision profiles which act as the major constraints on travel behaviour include : distributions of productions and attractions; distribution of zonal income levels; distributions of travel mode, travel time, travel cost, and employment/ education opportunities for each of the status group; and travel time and travel cost distributions by mode. The components of the model are activated in a specific order and with certain specific feedbacks. The model is so structured that the order of activities of the different components can be easily changed as per the framework of the study.

The simulation procedure first creates an individual. From the origin of the individual and income/education level, the status group is identified. The mode of travel is identified from the decision profile of travel generated for the relevant status group. The mode selected remains fixed during the decision cycle of the individual's destination choice. If it so happens that the destination could not be

assigned, then a new attempt on the selection of the mode is made. The travel time range for the individual is selected from the travel time decision profile of the relevant status group. Using the decision profile of employment/education opportunities, a set of destination zones that have the potential to attract, and can be reached within the assigned travel time range are identified. When more than one destination zones are identified, a travel cost range is randomly drawn from the appropriate decision profile. Out of the already identified set of destinations, only those which can be reached within the selected travel cost range are sorted, and the one with minimum travel cost is selected provided it has already not been assigned the desired trip ends.

The empirical validation of the simulation model is attempted by comparing the model outputs with the observed household survey data. Each of the submodels is validated by appropriate measures of effectiveness, so as to ensure that they have been assembled into a realistic structure . The results of the validation procedure indicate the capability of the model to simulate the decision process of the individual in a complex multimodal situation. The algorithm is quite sensitive and also flexible enough to incorporate alternative decision processes.

Programming Model for Combined Trip Distribution and Assignment

An attempt is also made in this study to derive models that combine trip distribution and traffic assignment into a single stage. This model optimizes the system by minimizing the total cost of the system as a whole. The main components

of the model are entropy type trip distribution submodel for each mode, and a system optimized assignment submodel. The combined model is solved by using the decomposition principle. All the constraints of the gravity submodel of the distribution problem are linear except the entropy constraint which is convex in the region defined by linear constraints. The gravity submodel is calibrated by making use of Davidon-Fletcher-Powell (DFP) method of constraint optimization.

The gravity submodel through an iterative process evolves the distribution matrix for each mode within the framework of the constraints of the given trip generation data, the generalized cost, and existing extent of entropy. The distribution matrix thus obtained is the input to the assignment submodel. The assignment submodel minimizes the total travel cost through system optimization for each mode considering the link costs over the network. The dual costs obtained from the assignment submodel replace the generalized costs in gravity submodel as feedback so as to get the revised trip matrices. These two submodels stop interacting when no further change in objective function is possible.

The assignment model requires alternative paths between each O-D pair as an input. In this study only two alternative paths between each O-D pair are considered which include the shortest travel time path and the one with next higher travel time. The capacity of each of the links also act as a constraint while assigning the trips. The model solution involves large magnitude of core capacity because of the requirements of work

vector, the coefficient matrices and resource vector inbuilt in the linear programming. The available DEC-1090 system is not capable of handling the problems of this magnitude. Hence it is decided to consider atmost 100 significant O-D pairs for assignment by the model. A strategy is planned for identifying O-D pairs having significant number of interzonal transfers. The assignment for these significant pairs is done by the model. Transfers between other O-D pairs are assigned to the shortest paths. The comparisons of the observed and simulated frequencies of trips for each of the travel time groups indicate a close agreement. The simulated interzonal transfer values are also close to the observed ones. The link flows as obtained from the model are compared with link flows when trips are assigned only to shortest paths. The results indicate that even for the city of Kanpur, where not many alternative paths are available, flows on some of the links do get affected while optimizing the system costs.

6.2 CONCLUSIONS

(1) The zonal regression equations constructed for work and education trips are statistically quite reliable indicating the validity of the explanatory variables selected in this study. Zonal regression is simple and quite manageable. Aggregating the basic survey data, however, removes most of the essential variability and could give a misleading high impression of statistical accuracy. Further the zonal regression equations may not provide an explanation of trip productions at levels other than the zonal aggregated

data on which they are based. Testing the invariance of the model over time will judge the capability of the model for forecasting trip productions.

(2) Cross classification analysis leads to a better understanding and explanation of trip generation than obtained from zonally aggregated data. Since the classification system is independent of zone system, the catalogue of characteristic mean trip rates can be applied to any system of aggregation for which the mixture of household types is known or can be forecasted. Subdivision of households into many types, however, provides uneven cell sizes making some of the mean trip rates less reliable than the others.

(3) The household regression analysis has produced sufficiently encouraging results to be preferred over zonal regression and household classification. Unlike zonal regression, household regression takes full account of the variability of the trip data at household level and provides a better understanding of trip productions than gained from mean trip rates corresponding to category analysis. The approach is independent of the zone system and the resulting equations are readily applicable to any system of areal units.

(4) The study attempts to explore the use of models based on non-compensatory choice processes in modal choice. These models appear to have a bright future to represent the mode choice decision processes. The results obtained

from maximin, and maximax models for the city of Kanpur confirm that the hypotheses drawn from non-compensatory models provide a sufficiently accurate prediction of mode choice decision processes. However the predominant-attribute model as formulated in this study does not provide reliable predictions. It is necessary that a broad range of choice models be tested before adopting one for explaining the mode choice behaviour.

(5) Simulation model is capable of simulating the decision process of the individuals in a complex multimodal situation. The model simulates the modal split and trip distribution as an interconnected decision process. The model components are activated in a specific order and with certain feedbacks. The structure of the model is such that the order of activities of different components can be easily changed if required as per the framework of the study. This has been attempted in a limited way for the case study. This flexibility of the model helps to evaluate the alternative decision processes.

The empirical validation of the various component submodels show that the model system is a sufficiently good replication of the complex multimodal situations. The model structure need to be extended to incorporate the traffic assignment aspect also.

(6) The programming model combines trip distribution and traffic assignment into a single stage and optimizes the system by minimizing the system cost. The algorithms evolved have been successfully applied to a large sizes network and are computationally efficient. The procedure of constrained optimization adopted for the calibration of gravity submodel is definitely better than the other empirical procedures. The strategy evolved for identifying the significant node pairs and having a combined distribution and assignment for these pairs first, help in solving large size networks.

The assignment submodel is capable of incorporating a number of alternative paths between individual node pairs. But the cities in developing countries have limited number of links in the network. The procedure adopted in this study of considering only two alternative paths in a particular iteration is computationally quite efficient without any significant deviations from the optimum values. The assignment algorithm is quite sensitive to the link costs and the link capacities. A reliable estimation of these measures is of great importance.

(7) The greatest contribution of the study is that the algorithms evolved are practical for problems encountered in cities of developing countries. Further the validity of the various models is clearly established. To enhance the achievements of the study, the models may be further refined and tested for a number of cities.

REFERENCES

Akcelik, R., 'A Graphical Explanation of the Two Principles and Two Techniques of Traffic Assignment', Transportation Research, Vol.13A, 1979.

Ashford, N. and Holloway, F., 'Time Stability of Zonal Trip Production Models', Proceeding of A.S.C.E.(Transp. Engg. Journal), 98, TE4, 1972.

Ashish Sen, and Johnson, Chris., 'A Simple Method for Analysis of Travel Data', Trans. Research, Vol. 11, June, 1977.

Beckmann, M., 'On the Metaphysical Foundations of Traffic Theory: Entropy Revisited', Traffic Flow and Transportation Proceedings, Fifth International Symposium on the Theory of Traffic Flow and Transportation, Elsevier, New York, 1971.

Ben-Akiva, M. , 'Structure of Passenger Travel Demand Models', Transportation Research Record 526, Transportation Research Board, Washington, D.C., 1974.

Ben-Akiva, M., and Adler, T., 'Social Recreation Joint Choice Model,' Working Paper No.8, Centre for Transportation Studies, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1975.

Ben-Akiva, M., Lerman, Steven, R., and Manheim, Marvin, L., 'Disaggregate Models: An Overview of Some Recent Results and Practical Applications', Transportation Models, PTRC Annual Meeting, University of Warwick, England, July, 1976.

Binder, Robert .H., 'Major Issues in Travel Demand Forecasting Brand, L. and Marwin, L. Manheim (eds), Urban Travel Demand Forecasting', Special Report 143, Washington, D.C. Highway Research Board, 1973.

Blundon, W.R., 'Some Applications of Linear Programming to Transportation and Traffic Problems', Institute of Traffic and Transportation Engineering, University of California at Berkeley, 1956.

Bouchard, Richard.J., 'Relevance of Planning Techniques to Decision-Making', Brand, Daniel L. and Marwin L. Manheim (eds), Urban Travel Demand Forecosting, Special Report 143, Washington, D.C., Highway Research Board, 1973.

Brokke, G.E., 'Evaluating Trip Forecasting Methods with an Electronic Computer', Bulletin No. 203, Highway Research Board, Washington, D.C., 1958.

Bruton, Michael. J., 'Introduction to Transportation Planning', Hutchinson of London, 1975.

Bruynooghe, M., Gilbert, A. and Sakarovitch,M., 'Methods of Effecting Transportation and Traffic', Proceedings of the Fourth Symposium on the Theory of Traffic Flow, Kallsruhe, 1968.

Bruynooghe, M., 'Combined Distribution and Assignment of Traffic', Report from Institute of Transport Research, I.R.T., Paris, 1969.

Bruzelius, Nils,,'The Value of Travel Time', Theory and Measurement, Croom Helm, London, 1979.

Bruzelius, Nils.,'Accessibility, Entropy and the Distribution and Assignment of Traffic-A Comment', Transportation Research-B, Vol.15B, 1981.

Bureau of Public Roads, 'Calibrating and Testing a Gravity Model for any Sized Urban Area', U.S. Department of Commerce, Washington, D.C., 1965.

Bulgren, William. G., 'Discrete System Simulation', Prentice Hall, Inc., Englewood Cliff's New Jersy,7632, 1982.

Burns, L.D., Golob, T.F., and Nicholaidis, G.C.,' A Theory of Urban Households Automobile Ownership Decisions', Transportation Research Record 569, 1976.

Cambridge Systematics, Inc., 'Introduction to Urban Travel Demand Forecasting', Urban Mass Transport Administration, Department of Transportation, 1974.

Chari, S.R., 'Estimation of Home Based Trip Productions by Regression Techniques', First Conference of the Road Engineering Association of Asia and Australia,Bangkok, Feb. 1976.

Charles River Associates, Inc., 'A Disaggregated Behavio Model of Urban Travel Demand', Federal Highway Administration, U.S. Department of Transportation,1972.

Chatterjee,A.,and Khasnabis,S.,'Category Models - A Case for Factor Analysis', Traffic Engineering, Vol.44, No.1, Oct. 1973.

Chicago Area Transportation Study, Final Report (in 3 parts), Chicago, 111, 1959, 1960 and 1962.

Christopher, Doubleday.,'Some Studies of the Temporal Stability of Person Trip General Models', Transp. Research, Vol.111, August, 1977.

Cochrane, R.A., 'A Possible Economic Basis for the Gravity Model', Journal of Transport Economics and Policy 9, 1975.

Coombs, C.H., 'A Theory of Data', Wiley, New York, 1964.

Dafermos, S.C., and Sparrow, F.T., 'The Traffic Assignment Problem for a General Network', Journal of Research, Natural Bureau of Standards 73B, 1969.

Dafermos, S.C., 'An Extended Traffic Assignment Model with Applications to Two-way Traffic', Transportation Science 5, 1971.

Dafermos, S., 'Integrated Equilibrium Flow Models for Transportation Planning', International Symposium on Traffic Equilibrium Models, Nov. 1974, Montreal, Canada, 1974.

Daganzo, C.F., 'On the Traffic Assignment Problem with flow Dependent Costs I and II', Transportation Research II, 1977 a,b.

Daganzo, C. F., 'An Equilibrium Algorithm for the Spatial Aggregation Problem of Traffic Assignment', Transportation Research, Vol.14B, No.3, 1980.

Detroit Metropolitan Area Traffic Study, Final Report, Detroit, Michigan, 1955.

Douglas, A.A., and Lewis, R.J., 'Trip Generation Techniques Household Least Square Regression Analysis', Traffic Engineering and Control, Jan. 1971.

Douglas, A.A., 'Home-Based Trip End Models - Comparison Between Category Analysis and Regression Analysis Procedure', Transportation, 1973.

Downes, J.D., 'Temporal Stability and Forecasting Ability of Trip Generation Models in Reading', Transportation Models, Summer Annual Meeting, University of Warwick, England, July,1976.

Einhorn, H.J., 'The Use of Nonlinear Non-compensatory Models in Decision Making', Psych. Bull.73, 1970.

Erlander, Sven.,'Accessibility, Entropy and the Distribution and Assignment of Traffic', Transportation Research, Vol.11, 1977.

Evans, A.W., 'Some Properties of Trip Distribution Models', Transportation Research. 4(1), 1970.

Evans, S.P., 'Some Application of Optimization Theory in Transport Planning', Dissertation Research Group in Traffic Studies, University College, London, 1973.

Evans, S.P., 'Derivation and Analysis of Some Models for Combining Trip Distribution and Assignment', Transportation Research, Vol.10, 1976.

Fleet, C., Stowers, J., and Swerdloff, C., 'Household Trip Production, Results of a Nation Wide Survey', Highway Planning Technical Report No.2, U.S. Bureau of Public Roads, Aug., 1965.

Fleet, C.R., and Robertson, S.R., 'Trip Generation in the Transportation Planning Process', Highway Research Record No. 240, Washington, D.C., 1968.

Florian, M., Nguyen, S., and Ferland, J., 'On the Combined Distribution-Assignment of Traffic', Publication No.8 Centre for Transportation Research, University of Montreal, 1974.

Florian, M., Nguyen, S., and Ferland, J., 'On the Combined Distribution and Assignment of Traffic', Transportation Science 9, 1975.

Florian, M., 'Traffic Equilibrium Methods', Lecture Notes in Economics and Mathematical Systems, Vol.118, Springer-Verlag, Berlin, New York, 1976.

Florian, M., and Nguyen, S., 'An Application and Validation of Equilibrium Trip Assignment Method', Transportation Science, 10, November, 1976.

Florian, M., and Nguyen, S., 'A Combined Trip Distribution, Model Split and Trip Assignment Model', University of Montreal Publication No. 34, 1977.

Foerster, J.F., 'Mode Choice Decision Process Models: A Comparison of Compensatory and Non-compensatory Structures', Transportation Research, Vol.13A, Pergamon Press Ltd., 1979.

Fox, Richard. L., 'Optimization Methods for Engineering Design', Addison-Wesley Publishing Company, Reading, Massachusetts, 1971.

Frank, M., and Wolfe, P., 'An Algorithm for Quadratic Programming', Naval Research Logistics Quarterly, 3, 1956.

Wrater, T.J., 'Forecasting Distribution of Interzonal Vehicular Trips by Successive Approximations', Proceedings,Highways Research Board, Washington, D.C., Vol.33, 1954.

Green, P.E., and Wind, Y., 'Multiattribute Decisions in Marketing', Dryden Press, Hinsdale, Illinois, 1973.

Hansen, Stein., 'Utility, Accessibility and Entropy in Spatial Modelling', Swedish Journal of Economics 74, 1972.

Hughes, G.D., 'Buyer/Consumer Information Processing: An Overview of Where Researchers have been and Where They Should be Going', University of North Carolina Press, Chapel Hill, N.C., 1974a.

Hughes, G.D., 'Trends in the Development of Multi-attribute Models of Central Information Processing (Edited by G.D. Hughes and M.L.Ray)', University of North California Press, Chapel Hill, N.C. 1974b.

Hutchinson, B.G., 'Principles of Urban Transport Systems Planning', McGraw-Hill Book Company, New York,1974.

Jefferson, T.R., and Scott, C.H., 'The Analysis of Entropy Models with Equality and Inequality Constraints', Transportation Research,B, Vol.13B, 1979.

Jones, J. Morgan, 'Introduction to Decision Theory', Richard D, Drwin, Inc., Homewood, Illinois 60430, 1977.

Kannel, E.J., 'A Disaggregate Analysis of Urban Travel Behaviour', Joint Highway Research Project, No.17, Purdue University, Indiana, July 1972.

Kassoff, H., and Deutschman, H.P., 'Trip Generation A Critical Appraisal', Highway Research Record No.297, 1969.

Kleinmuntz, B., 'Formal Representation of Human Judgement', Wiley, New York, 1968.

Kreibich, Volker., 'Modelling Car Availability, Modal Split and Trip Distribution by Monte-Carlo Simulation: A Short Way to Integrated Models', Transportation,8,1979.

Lansing, J.B., 'The Effects of Migration and Personel Effectiveness of Long Distance Travel', Transportation Research, Vol.2, No.4, Dec., 1968.

Lave, C.A., 'A Behavioral Approach to Modal Split Forecasting', Transportation Research,3, 1967.

Leblanc, L.J., 'Mathematical Programming Algorithms for Large Scale Network Equilibrium and Network Design Problems', Ph.D. Dissertation, Department of Industrial Engineering and Management Sciences, Northwestern University, 1973.

Leblanc, L.J., Morlok, E.K., and Pierskalla, W.P., 'An Efficient Approach to Solving the Road Network Equilibrium Traffic Assignment Problem, Transportation Research, 9, 1975.

Lerman, Steven. R., 'A Disaggregate Behavioral Model of Urban Mobility Decisions,' Ph.D. Dissertation, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1975.

Lerman, Steven. R., and Ben-Akiva, M., 'A Disaggregate Behavioural Model of Auto Ownership', Transportation Research Record 569, Transportation Research Board, Washington, D.C., 1975.

Lerman, S.R., and Moshe, B.A., 'A Disaggregate Behavioral Model of Automobile Ownership', Transportation Research Record 569, 1976.

Leventhal, T.L., Nemhauser, G.L., and Trotter, L.E., 'A Column Generation Algorithm for Optimal Traffic Assignment', Transportation Science, 7, 1973.

Lisco, T.E., 'The Value of Commuters Travel Times: A Study in Urban Transportation', Doctoral Dissertation, University of Chicago, Chicago, Illinois, 1967.

MacCrimmon, K.R., 'Decision Making Among Multiple Attribute Alternatives: A Survey and Consolidated Approach', Rand Memorandum RM-4823-ARPA, Santa Monica, 1968.

Marksjo, B., 'Gravity Model in Traffic Planning', PT 19702, Stockholm, 1970.

Martin, B.V., Memmott, F.W., and Bone, A.J., 'Principles and Techniques of Predicting Future Demand for Urban Area Transportation', MIT, Massachusetts, 1961.

McFadden, Danial. L., 'The Theory and Practice of Disaggregate Demand Forecasting for Various Modes of Transportation', Emerging Transportation Planning Methods, Department of Transportation, Washington, D.C., 1978.

McCarthy, Gerald, M., 'Multiple Regression Analysis of Household Trip Generation', Highway Research Record, No.297, 1969.

McGillavary, Robert. G., 'Binary Choice of Urban Transport Mode in the San Francisco Bay Region', Econometrica, Vol.40, No.5, September, 1972.

Murchland, J.D., 'Some Remarks on the Gravity Model of Traffic Distribution and an Equivalent Maximization Formulation', London School of Economics, 1966.

Murchland, J.D., 'Road Network Traffic Distribution in Equilibrium', Paper Presented at the Conference on Mathematical Methods in the Economics Sciences, Oberwolfach, 1969.

Neter, John., and Wasserman, William., 'Applied Linear Statistical Models', Richard D.Irwin, Inc., Home Wood, Illinois, 1974.

Neuberger, H.L.T., 'User Benefit in the Evaluation of Transport and Land Use Plans', Journal of Transport Economics and Policy, 5, 1973.

Nguyen, S., 'An Algorithm for the Traffic Assignment Problem', Transportation Science 8, 1974.

Nie, Norman. H., and Partners., 'Statistical Package for the Social Sciences', Second Edition, McGraw Hill Book Company, 1977.

Nijkamp, P., 'Reflections on Gravity and Entropy Models', Regional Science and Urban Economics, 5, 1975.

Oi, W.Y., and Shuldiner, P.W., 'An Analysis of Urban Travel Demands', North Western University Press, Evanston, Illinois, 1962.

Parsons, Robert., 'Statistical Analysis: A Decision Making Approach', Harper and Row, Publishers, New York, 1974.

Patnaik, S.B., 'Programming for O timal Design of an Urban Bus Transit. Network System', Ph.D. Thesis, I.I.T. Kanpur, India, 1983.

Paul, O. Roberts, Jr., 'Disaggregate Demand Forecasting: Theoretical Tantalizer or Practical Problem Solver', Emerging Transportation Planning Methods, Department of Transportation, Washington, D.C., 1978.

Peat., Marwick., Michell and Company., 'Implementation of the n-Dimensional Logit Model', Prepared for the Comprehensive Planning Organization, San Diego County, California, May, 1973.

Potts, R.B., and Oliver, M.R., 'Flows in Transportation Network', Academic Press, New York, 1972.

Quaramby, D.A., 'Choice of Travel Mode for the Journey to Work: Some Findings', Journal of Transport Economic Policy, 1967.

Ruiter, E.R., 'Implementation of Operational Network Equilibrium Procedures' Transportation Research Record 491, 1974.

Russ, F.A., 'Evaluation Process Models: Objective and Subjective Comparisons' 19th Int. Tims Meeting, Houston, Tx, 1972.

Scheele, Shiv., 'A Mathematical Programming Algorithm for Optimal Bus Frequencies', Linkoping University, Institute of Technology, Sweden, 1977.

Shuldiner, P.W., 'Trip Generation and Home', Highway Research Board Bulletin, 347, 1962.

Smith, R.L., and Cleaveland, D.E., 'Time Stability of Trip Generation and Pedestrian Modal Choice Models', Transportation Research Record, No. 569, 1976.

Steenbrink, P.A., 'Transport Network Optimization in the Dutch Integral Transportation Study, Transportation Research, 8, 1974.

Stopher, P.R., 'A Probability Model of Travel Mode Choice for the Journey to Work', Highway Research Record, 283, 1969.

Stopher, P.R., and Lisco, T.E., 'Modelling Travel Demand', A Disaggregate Behavioural Approach Issues and Applications', Transportation Research Form, Vol. 11, 1970.

Stopher, P.R., and Meyburg, A.H., 'Urban Transportation Modelling and Planning', D.C. Heath, Lexington, M.A., 1975.

Talvitie, A., 'Aggregate Travel Demand Analysis with Disaggregate and Aggregate Travel Demand Models', Transportation Research Form, Vol. 14, 1973.

Tardiff, T.J., 'Comparison of Effectiveness of Various Measures of Socio-Economic Status in Models of Transportation Behaviour', Transportation Research Record, No. 534, 1975.

Tardiff, T.J., 'Trip Generation as a Choice Process', Journal of Transportation Engineering, Vol. 103, March, 1977.

Taylor, W.C., 'Optimization of Traffic Flow Splits', Highway Research Record, No. 230, Highway Research Board, Washington, D.C., 1968.

Thomson, Michael. J., 'Toward Better Urban Transport Planning in Developing Countries', World Bank Staff Working Papers, Number 600, The World Bank, Washington, D.C., USA, 1983.

Tomlin, J.A., and Tomlin,S.G., 'Traffic Distribution and Entropy', Nature, Vol.220, 1968.

Tomlin, J.A., 'A Mathematical Programming Model for the Combined Distribution-Assignment of Traffic', Transportation Science, Vol.5, No.2, 1971.

Wardrop, J.G., 'Some Theoretical Aspects of Road Traffic Research', Proceedings, Institution of Civil Engineering, Part II, 1952.

Warner, S.L., Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice', North Western University Press, Evanston, IL, 1962.

Wigan, M.R., 'New Techniques for Transport Systems Analysis', Australian Road Research Board Special Report ,No. 10, 1977.

Wilson, A.G., 'A Statistical Theory of Spatial Distribution Models', Transportation Research, 1, 1967.

Wilson, A.G., 'Entropy in Urban and Regional Modelling', Pion Limited, 1970.

Wootton, H.J., and Pick,G.W., 'A Model of Trips Generated by Households', Journal of Transport Economics and Policy, May, 1967.

World Bank ,'Urban Transport', A Sector Policy Paper, May,1975.

World Bank , 'Kanpur Urban Development Project', Report No. 3504-IN, 1981.

APPENDIX I

DETAILS OF TRAFFIC ZONES WITH CORDONS NUMBERS FOR KANPUR CITY

Traffic zone no.	Name of the area	Cordon no.
(1)	(2)	(3)
01	National Sugar Institute, Bisekhpur Banzar, Azad Nagar, Kheora Kachhar, Bishan Kachhar, Lakhanpur, Akbarpur Kachhar, Baribazar Naubasta.	3
02	Nawabganj (A)	3
03	Nawabganj (B)	3
04	Tilak Nagar, Benajhavar, Swaroop Nagar	2
05	Arya Nagar	2
06	Gwal Toli, Khalsi Line	2
07	Bhairon Ghat, Soorganj	3
08	Parmat	3
09	Civil Lines	2
10	McRobertganj	2
11	Naresh Nagar	2
12	Ram Bag, Jawahar Nagar, Gandhi Nagar, Nehru Nagar, Ram Krishna Nagar, Ashok Nagar	1
13	Sisamau (A)	1
14	Chamanganj, Sisamau (B)	1
15	Talaq Mohal, Baconganj, Colonelganj(A)	1
16	Colonelganj (B)	1
17	Civil Lines	2

Contd...

APPENDIX I contd...

(1)	(2)	(3)
18	–	
19	Media Bazar, Naya Chowk, Faras Khana, Bisati Bazar, Boochar Khana, Moolganj	1
20	Anwarganj, Iftikarabad, Dalel Purwa(A), Faras Khana, Nazir Bag	1
21	Bans Mandi, Dalel Purwa(B)	1
22	Rai Purwa, Bhannana Purwa	1
23	Cooperganj, Laxmi Purwa	1
24	Collectorganj, Dhankutti, Ranjit Purwa, Coolie Bazar, Boochar Khana, Coperganj	1
25	Chatri Muhal, Feel Khana, Lathi Muhal, Purani Sabji Mandi, Chowk Patri, Hata Sawai Singh, Bangali Muhal, Mall Road, Narial Bazar, Chowk Sarafa, Chapra Muhal, Hatia	1
26	Civil Lines	2
27	Kursawa, Patkapur, Chatai Muhal, Peel Khana, Karachi Khana, Roti Godam, Sirki Muhal, Nachghar, Dal Mandi	1
28	Generalganj, Naudhara, Ramganj, Nayaganj	1
29	Harbansh Muhal, Gadariya Muhal, Kachiana Muhal, Daulatganj, Lokman Muhal, Dana Khori, Madhuri Muhal, Sutar Khana, Moti Muhal	1
30	Northern Railway Colony	3
31	Contonment (A)	3
32	Contonment (B)	3
33	Gangaganj, Deviganj, Tabedapur, Kamalpur, Safipur, Gewakhera, Kazikhera, Chandari, Pardawanpur Gaushala, Wazidpur	3
34	Pokharpur, Sarai Ramaku, Arzai Pardwanpur, Gangoopura, Wezadpur	3
35	Chakeri	3

contd....

APPENDIX I contd...

(1)	(2)	(3)
36	Pendi, Manega, Sekhapur, Jena, Kishanpur, Madarpur, Kulgaon, Khajuria	3
37	Roma, Chakeri (Village), Tikri Paigambarpur, Tikri, Chitara, Gorekha, Kishan Nagar	3
38	Bharpur Macharia, Lalpur, Todakpur	3
39	C.O.D.	3
40	Babupurwa Munshipurwa, Newhara	3
41	Kidwai Nagar, Babupurwa Colony, Azibganj Bagahi, Chandari Begampurwa	3
42	Juhi Hamirpur Road (A)	2
43	Juhi Khurd, Juhi Hamirpur Road (B)	2
44	Central Railway Colony	2
45	Zarib Chowki	1
46	Kaushalpuri, Darshanpurwa, Nasimabad, Krishna Nagar	1
47	Kamal Nagar, Shastri Nagar, Vijay Nagar	2
48	Sarojini Nagar, Fazalganj	2
49	Govind Nagar Gohi Colony, Govind Nagar	2
50	Hanspur, Bingawan, Pokharpur, Aira, Naubasta, Juhi Kala, Usmanpur, Chorepur, Zaroli, Kheharai, Vivekanand Nagar	3
51	Newaraiya Khera, Jamu, Panki	3
52	Shahpur, Sazanpur, Panki Gangaganj, Panki Katra	3
53	Armapur	3
54	Kakadeo	3
55	Bhinekhapur, Masanpur, Rawatpur (village) Daulatpur, Kalyanpur Kala, Kalyanpur Khurd, Mirzapur, Bhawanipur, Gambhirpur	3
56	I.I.T.	3

APPENDIX II

ALTERNATIVE PATHS BETWEEN IMPORTANT O-D PAIRS

O-D Pair	Total no. of nodes	Sequence of nodes
(1)	(2)	(3)
02-25	7	02 04 05 10 16 19 25
	8	02 04 05 11 10 16 18 25
02-46	4	02 54 11 46
	6	02 04 05 10 12 46
02-55	2	02 55
	3	02 01 55
03-04	2	03 04
	3	03 02 04
03-09	3	03 07 09
	4	03 07 06 09
03-12	5	03 04 05 10 12
	6	03 07 06 10 11 12
03-18	4	03 07 09 18
	5	03 07 09 16 18
03-19	5	03 07 09 18 19
	5	03 07 09 16 19
03-20	5	03 07 09 16 20
	6	03 07 09 16 15 20
03-25	6	03 07 09 18 19 25
	7	03 07 09 16 19 18 25
03-26	5	03 07 09 17 26
	5	03 07 08 17 26
03-28	7	03 07 09 18 19 25 28
	8	03 07 09 16 19 18 25 28
03-35	8	03 07 09 17 26 32 33 35
	8	03 07 08 17 26 33 34 35

Contd.....

(1)	(2)	(3)
03-51	5 5	03 02 54 53 51 03 04 54 53 51
03-56	4 4	03 02 01 56 03 02 55 56
04-05	2 3	04 05 04 06 05
04-06	3 5	04 05 06 04 11 05 10 06
04-08	5 6	04 05 06 07 08 04 11 05 06 09 08
04-09	4 6	04 05 06 09 04 11 05 06 07 09
04-12	4 3	04 05 10 12 04 11 12
04-15	5 4	04 05 10 16 15 04 11 12 15
04-18	5 6	04 05 10 16 18 04 11 05 06 09 18
04-25	6 11	04 05 10 16 19 25 04 11 05 06 10 15 16 20 19 18 25
04-28	7 13	04 05 10 16 19 25 28 04 11 05 06 10 15 16 20 19 18 25 27 28
04-29	8 9	04 05 06 09 17 26 27 29 04 11 05 10 16 19 25 28 29
04-35	9 9	04 05 06 09 17 26 32 33 35 04 06 10 16 18 26 32 33 35
04-46	5 3	04 05 10 12 46 04 11 46
04-47	3 5	04 54 47 04 02 54 53 47
04-48	6 4	04 05 10 12 46 48 04 11 46 48
05-12	3 3	05 10 12 05 11 12

Contd.....

APPENDIX II contd...

(1)	(2)	(3)
06-18	3	06 09 18
	4	06 10 16 18
06-25	5	06 09 18 19 25
	5	06 10 16 18 25
06-30	7	06 09 18 19 24 31 30
	9	06 09 16 15 14 22 45 44 30
06-35	7	06 09 17 26 32 33 35
	7	06 09 17 26 32 34 35
06-37	7	06 09 18 19 24 31 37
	7	06 10 16 19 24 31 37
06-55	5	06 05 04 02 55
	6	06 05 11 04 54 55
07-18	3	07 09 18
	4	07 09 16 18
07-25	5	07 09 18 19 25
	6	07 09 16 19 18 25
09-06	2	09 06
	3	09 07 06
09-15	3	09 16 15
	4	09 18 20 15
09-17	2	09 17
	3	09 08 17
09-19	3	09 18 19
	3	09 16 19
09-25	4	09 18 19 25
	5	09 16 19 18 25
09-26	3	09 17 26
	3	09 18 26
09-48	6	09 16 15 12 46 48
	5	09 10 12 46 48
10-18	3	10 16 18
	4	10 16 19 18
10-24	4	10 16 19 24
	5	10 16 18 19 24

Contd,..

APPENDIX II contd...

(1)	(2)	(3)
10-46	3	10 12 46
	3	10 11 46
11-04	3	11 05 04
	4	11 10 06 04
11-08	5	11 10 06 07 08
	5	11 10 16 09 08
11-09	4	11 10 16 09
	4	11 10 06 09
11-15	4	11 10 16 15
	3	11 12 15
11-25	5	11 10 16 19 25
	6	11 10 15 16 18 25
11-28	6	11 10 16 19 25 28
	10	11 10 15 16 20 19 18 25 27 28
12-01	6	12 10 05 04 02 01
	7	12 10 05 04 03 02 01
12-09	4	12 15 16 09
	4	12 10 16 09
12-15	2	12 15
	3	12 13 15
12-18	4	12 15 16 18
	4	12 10 16 18
12-19	4	12 15 20 19
	5.	12 13 14 20 19
12-23	5	12 13 14 21 23
	4	12 13 22 23
12-24	5	12 15 20 19 24
	6	12 13 14 20 19 24
12-25	5	12 15 20 19 25
	8	12 13 14 20 24 19 18 25
12-28	6	12 15 20 19 25 28
	9	12 13 14 20 19 18 25 27 28
12-31	6	12 13 14 21 23 31
	6	12 13 22 23 24 31

Contd....

(1)	(2)	(3)
12-46	2 3	12 46 12 13 46
12-47	2 3	12 47 12 46 47
12-48	3 4	12 46 48 12 13 46 48
12-49	4 5	12 45 44 49 12 13 45 44 49
13-08	6 6	13 14 15 16 09 08 13 14 20 16 09 08
13-25	5 6	13 14 20 19 25 13 15 20 19 18 25
13-34	7 8	13 14 21 23 31 33 34 13 22 23 24 31 32 33 34
14-08	5 5	14 15 16 09 08 14 20 16 09 08
14-09	4 4	14 15 16 09 14 20 16 09
14-15	2 3	14 15 14 20 15
14-25	4 4	14 20 19 25 14 15 20 25
14-31	4 5	14 21 23 31 14 22 23 24 31
15-25	4 5	15 20 19 25 15 16 19 18 25
15-35	7 10	15 20 21 23 31 33 35 15 14 21 23 24 31 32 33 37 35
17-06	3 5	17 09 06 17 08 09 07 06
17-12	5 5	17 09 16 15 12 17 09 16 10 12
17-18	2 3	17 18 17 26 18

Contd.....

APPENDIX II contd...

(1)	(2)	(3)
17-25	4 3	17 18 19 25 17 26 25
17-28	4 5	17 26 27 28 17 18 19 25 28
17-31	4 5	17 26 27 31 17 18 19 24 31
17-45	7 7	17 09 16 15 14 22 25 17 09 18 20 21 23 45
17-42	9 9	17 09 16 15 14 22 45 44 42 17 08 19 20 21 23 45 44 42
19-25	2 3	19 25 19 18 25
19-47	6 6	19 20 14 13 46 47 19 20 15 12 46 47
20-18	3 2	20 19 18 20 18
20-23	3 4	20 21 23 20 14 21 23
20-35	6 8	20 21 23 31 33 35 20 19 24 31 32 33 37 35
21-12	4 4	21 14 13 12 21 22 13 12
21-15	2 3	21 15 21 20 15
21-25	4 4	21 20 19 25 21 14 20 25
21-28	3 5	21 24 28 21 20 19 25 28
22-15	3 4	22 14 15 22 21 20 15
22-28	4 6	22 21 24 28 22 21 20 19 25 28
23-18	4 5	23 21 20 18 23 21 20 19 18

Contd....

APPENDIX II contd...

(1)	(2)	(3)
24-08	5 5	24 19 18 09 08 24 19 16 09 08
24-18	3 4	24 19 18 24 19 16 18
24-25	3 6	24 19 25 24 21 20 19 18 25
24-28	2 4	24 28 24 19 25 28
24-35	4 5	24 31 33 35 24 31 32 34 35
25-08	5 6	25 19 18 09 08 25 18 19 16 09 08
25-28	2 3	25 28 25 27 28
25-35	6 7	25 19 24 31 33 35 25 28 31 32 33 37 35
25-48	8 7	25 19 20 21 22 45 44 48 25 19 20 15 12 46 48
26-25	4 3	26 18 19 25 26 18 25
27-06	5 6	27 26 17 09 06 27 26 17 09 07 06
27-08	4 5	27 26 17 08 27 26 17 09 08
27-09	4 4	27 26 17 09 27 26 18 09
27-12	6 6	27 26 18 16 15 12 27 26 18 16 10 12
27-18	3 4	27 26 18 27 25 19 18
27-19	3 4	27 25 19 27 26 18 19
27-25	2 5	27 25 27 26 18 19 25

Contd....

APPENDIX II contd...

(1)	(2)	(3)
31-32	2	31 32
	3	31 27 32
31-45	5	31 23 21 22 45
	6	31 24 23 22 46 45
31-47	6	31 23 21 22 46 47
	6	31 24 23 22 46 47
31-48	4	31 30 44 48
	7	31 23 21 22 45 44 48
31-25	5	32 26 18 19 25
	3	32 27 25
33-19	4	33 31 24 19
	5	33 32 26 18 19
33-25	5	33 31 24 19 25
	4	33 31 28 25
41-04	8	41 43 44 45 12 10 05 04
	11	41 42 43 44 48 46 13 12 10 05 04
41-08	10	41 43 44 45 22 14 15 16 09 08
	11	41 42 43 44 48 45 22 21 20 16 09 08
41-09	9	41 43 44 45 22 14 15 16 09
	12	41 42 43 44 48 46 22 14 21 20 16 09
41-12	5	41 43 44 45 12
	7	41 43 44 48 45 13 12
41-14	6	41 43 44 45 22 14
	7	41 42 44 48 45 13 14
41-15	7	41 43 44 45 22 14 15
	9	41 42 44 48 45 22 21 20 15
41-18	9	41 43 44 45 22 14 15 16 18
	12	41 42 43 44 48 46 22 14 21 20 19 18
41-19	8	41 43 44 45 22 21 20 09
	10	41 42 43 44 48 46 22 12 20 19
41-20	7	41 43 44 45 22 21 20
	8	41 43 44 48 45 22 14 20
41-23	6	41 43 44 45 22 23
	6	41 43 44 48 45 23

Contd....

APPENDIX II contd...

(1)	(2)	(3)
41-24	4 9	41 30 31 24 41 43 42 44 45 22 14 21 24
41-25	9 11	41 43 44 45 22 21 20 19 25 41 43 44 48 45 23 22 14 20 18 25
41-28	4 5	41 30 31 28 41 30 31 29 28
41-30	2 3	41 30 41 39 30
41-35	5 4	41 30 31 33 35 41 39 37 35
41-39	2 3	41 39 41 30 39
41-42	2 3	41 42 41 43 42
41-46	5 6	41 43 44 45 46 41 42 44 48 45 46
41-47	5 5	41 43 44 48 47 41 42 44 48 47
41-48	4 4	41 43 44 48 41 42 44 48
41-49	3 5	41 43 49 41 42 43 44 49
42-09	8 10	42 44 45 22 14 15 16 09 42 44 48 46 22 14 21 20 16 09
42-15	6 8	42 44 45 22 14 15 42 44 48 45 22 21 20 15
42-19	7 8	42 44 45 22 21 20 19 42 44 48 46 22 14 20 19
42-24	6 7	42 44 45 22 21 24 42 44 48 46 22 23 24
42-25	8 10	42 44 45 22 21 20 19 25 42 44 48 46 22 14 20 19 18 25
42-26	9 11	42 44 45 22 14 15 16 18 26 42 44 48 46 22 14 15 16 09 17 26

Contd....

APPENDIX II contd....

(1)	(2)	(3)
42-31	3	42 30 31
	4	42 41 30 31
42-46	4	42 44 45 46
	4	42 44 48 46
42-49	3	42 44 49
	3	42 43 49
43-12	4	43 44 45 12
	7	43 42 44 48 45 13 12
43-17	9	43 44 45 22 14 15 16 09 17
	11	43 42 44 48 45 22 21 20 16 09 17
43-28	7	43 44 45 22 21 24 28
	11	43 44 48 45 23 22 21 20 19 25 28
43-31	4	43 44 30 31
	4	43 41 30 31
43-37	4	43 41 39 37
	5	43 42 41 38 17
43-39	3	43 41 39
	4	43 42 41 39
43-41	2	43 41
	3	43 44 45 12
43-42	2	43 42
	3	43 44 42
43-48	3	43 44 48
	6	43 42 44 45 46 48
43-49	2	43 49
	3	43 44 49
44-46	3	44 45 46
	3	44 48 46
45-25	6	45 22 21 20 19 25
	5	45 22 14 23 25
46-08	6	46 12 10 06 07 08
	7	46 47 12 15 16 09 08
46-15	3	46 12 15
	3	46 13 15

Contd......

APPENDIX II contd...

(1)	(2)	(3)
46-18	5 5	46 12 15 16 18 46 12 15 20 19 25
46-24	5 4	46 45 22 21 24 46 22 23 24
46-25	6 8	46 12 15 20 19 25 46 13 14 20 24 19 18 25
46-35	7 7	46 22 21 23 31 33 35 46 22 23 24 31 37 35
46-47	2 3	46 47 46 48 47
46-48	2 4	46 48 46 45 44 48
46-49	4 3	46 45 44 49 46 48 49
46-51	4 4	46 48 53 51 46 48 49 51
47-18	6 5	47 46 12 15 16 18 47 12 15 16 18
47-23	5 4	47 46 22 21 23 47 46 22 23
47-24	6 6	47 46 45 22 21 24 47 12 15 20 19 24
47-25	7 8	47 46 12 15 20 19 25 47 54 11 10 16 19 18 25
47-28	8 9	47 46 12 15 20 19 25 28 47 46 13 14 20 19 18 25 28
47-33	7 8	47 46 22 21 23 31 33 47 46 22 23 24 31 32 33
47-34	8 9	47 46 22 21 23 31 33 34 47 46 22 23 24 31 32 33 34
47-35	8 10	47 46 22 21 23 31 33 35 47 46 22 23 24 31 32 33 37 35
47-45	3 4	47 46 45 47 48 44 45

Contd.....

(1)	(2)	(3)
47-48	2	47 48
	3	47 46 48
47-51	3	47 53 51
	4	47 54 53 51
48-08	7	48 46 12 15 16 09 08
	6	48 46 12 10 09 08
48-18	6	48 46 12 15 16 18
	6	48 46 12 10 16 18
48-28	7	48 44 45 22 21 24 28
	9	48 44 45 23 21 20 19 25 28
49-17	9	49 44 45 22 14 15 16 09 17
	11	49 44 48 45 23 22 21 20 16 09 17
49-18	8	49 44 45 22 14 15 16 18
	10	49 44 48 45 23 22 21 20 19 18
49-19	7	49 44 45 22 21 20 19
	8	49 44 48 46 22 14 20 19
49-23	6	49 44 45 22 21 23
	6	49 44 48 46 22 23
49-24	6	49 44 45 22 21 24
	7	49 48 46 13 22 23 24
49-25	8	49 44 45 22 21 20 19 25
	6	49 44 30 31 28 25
49-26	9	49 44 45 22 14 15 16 18 26
	11	49 44 48 46 22 14 15 16 09 17 26
49-27	9	49 44 45 22 21 20 19 25 27
	10	49 44 45 23 14 20 19 18 25 27
49-28	7	49 44 45 22 21 24 28
	12	49 44 48 46 22 14 21 20 19 18 25 28
49-29	5	49 44 30 31 29
	9	49 44 45 22 14 21 23 31 29
49-35	6	49 44 30 31 33 35
	12	49 44 45 22 14 21 23 31 32 33 37 35
49-47	4	49 44 48 47
	6	49 44 45 22 46 47
49-48	2	49 48
	3	49 44 48

APPENDIX II contd...

(1)	(2)	(3)
49-51	2 3	49 51 49 53 51
50-41	2 3	50 41 50 43 41
53-46	3 3	53 48 46 53 47 46
53-47	2 3	53 47 53 54 47
54-04	2 3	54 04 54 02 04
54-12	3 3	54 02 04 54 11 12
54-19	5 6	54 11 10 16 19 54 11 10 16 18 19
54-25	6 9	54 11 10 16 19 25 54 11 10 15 16 20 19 18 25
54-28	7 11	54 11 10 16 19 25 28 54 11 10 15 16 20 19 18 25 27 28
54-46	3 3	54 11 46 54 47 46
54-47	2 3	54 47 54 53 47
54-49	5 6	54 47 48 44 49 54 11 46 45 44 49
55-09	6 5	55 02 04 05 06 09 55 02 03 07 09
55-25	7 7	55 54 11 10 16 19 25 55 54 11 10 16 18 25
55-46	4 4	55 53 48 46 55 54 11 46
55-48	3 4	55 53 48 55 54 47 48
55-49	3 5	55 53 49 55 53 48 44 49
56-55	2 3	56 55 56 01 55